IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 18, Issue 2 Ser. 7 (February. 2019), PP 84-88
www.iosrjournals.org

Pre-operative Assessment of Acute Appendicitis Severity with Reference to a Structured Intra-Operative Grading System

Jacqueline Bakhit[1], Julian Choi[1,2], Lucy Lu[1], Carolyn Chew[1], Anh Vu[1], Steven Chan[1,2]

[1] *Department of Surgery, Western Health, Melbourne, Australia*
[2] *Melbourne Medical School, Western Health, Melbourne, Australia*
Corresponding Author: Jacqueline Bakhit

Abstract:
Background: *The previously published Sunshine Appendicitis Grading System (SAGS) score was developed to standardize the intra-operative severity of acute appendicitis. The authors proposed a novel method of pre-operative assessment using the SAGS score. The primary objective of this prospective study was to evaluate the ability of clinicians to assess the severity of acute appendicitis with reference to a standardized grading system.*
Methods: *One hundred and thirty-one patients who underwent laparoscopy for suspected acute appendicitis at Sunshine Hospital were assessed. Based on clinical findings, patients were scored for suspected appendicitis severity at the time of clinical assessment by clinicians of the surgical unit. They were scored again following laparoscopy based on the SAGS score. Log-linear modelling of ordinal agreement data was used to find the model of best-fit.*
Results*: The pre-operative and intra-operative SAGS score agreement index, weighted kappa (K_ω) was 0.62 (95% confidence interval 0.51 to 0.74). The best-fit model of diagonal agreement plus linear-by-linear association showed no systematic bias (deviance χ^2 7.56, P = 0.91).*
Conclusion: *Pre-operative SAGS score assessment demonstrated a significant agreement with intra-operative findings and is a useful aid in decision-making under uncertainties in cases of suspected appendicitis.*
Keywords: *appendicitis, laparoscopic appendectomy, Sunshine Appendicitis Grading System*

Date of Submission: 01-02-2019 Date of acceptance: 16-02-2019

I. Introduction

Acute appendicitis is a common surgical emergency. The diagnosis of appendicitis remains largely clinical and is dependent on optimal skills in history taking and eliciting physical signs. In some patients, judicious use of laboratory markers and radiological imaging may provide additional evidence to support a clinical diagnosis [1,2].

Multiple clinical scoring systems exist for the diagnosis of appendicitis. The Alvarado score is the most well-known scoring system with the goal of reducing negative appendectomy rates without increasing the risk of perforation [3]. There has been more recent introduction of the Appendicitis Inflammatory Response (AIR) scoring system and the APPEND score [4.5]. Both the AIR score and the APPEND score can stratify patients into low to high-risk of having appendicitis, but have not been shown to effectively gauge the severity of appendicitis with reference to a standardized intra-operative grading system.

The Sunshine Appendicitis Grading System (SAGS) score (Table 1) has been shown to simply and accurately grade appendicitis severity at the time of operation, predict the likelihood of post-operative collections and has reduced the ambiguity of the definition of "complicated" appendicitis at Western Health [6]. The primary objective of this study was to evaluate the ability of clinicians to pre-operatively grade the severity of appendicitis with referenced to the SAGS score.

Table 1: SAGS Score

SAGS Score	Intra-operative Findings
0	No appendicitis
1	Simple appendicitis (any of the following): i. Injected appendix, ii. Thickened appendix, iii. Serous free fluid
2	Purulent appendicitis (any of the following): i. Pus localized to right iliac fossa ii. Right paracolic gutter iii. Pelvis
3	Purulent appendicitis with 4 quadrant contamination
4	Perforated appendix (any of the following): i. Free fecalith, feces ii. Fecal staining iii. Visible hole in appendix

II. Materials & Methods

This was a single-institution prospective study. Patients presenting to the Sunshine Hospital Emergency Department were recruited between August 2017 and August 2018. Ethics approval was obtained. Inclusion criteria included patients ≥10 years of age who presented with suspected acute appendicitis that proceeded to have an operation. Patients were excluded if they had a Computed Tomography (CT) or ultrasound scan assessing the appendix prior to pre-operative assessment or if they did not proceed to laparoscopy.
Prior to commencement of the study, the clinicians were educated on the SAGS scoring system (Table 1).

Information was collected on a structured data collection form. Data included patient demographics, the presence or absence of the symptoms and signs of appendicitis, white-cell-count (WCC), neutrophil count and C-reactive-protein (CRP). A pre-operative SAGS score was then assigned to the patient at the time of assessment by a clinician (residents, registrars, fellows and consultants) on the surgical unit. The decision to proceed to laparoscopy was consultant-led. Patients were assigned an intra-operative SAGS score during laparoscopy by the operating surgeon.

2.1 Statistical Analysis

For pre- and intra-operative ratings, a minimum sample size of 80 patients was calculated to achieve a power of 80% (2-sided α 0.05) to detect a weighted kappa (K_ω) agreement index of 0.8. Doubly ordered categorical data were tested for row and column independence using linear-by-linear association test. To account for systematic bias amongst raters where one gives consistently higher (or lower) ratings than the other, log-linear modelling was used to examine the data for *lack-of-fit* between observed and expected frequencies and to find the model of *best-fit*.[7-9] For log-linear modelling the null hypothesis (H_0): the model is the correct model of best-fit relative to the baseline model of perfect fit and $P > 0.05$ signifies acceptance of the model;[8.9] otherwise $P < 0.05$ (2-sided) was considered statistically significant. All analyses were undertaken with StatXact v9 (Cytel Inc. Cambridge MA, USA) and Stata v14 (StataCorp, College Station, TX, USA).

III. Results

One hundred and thirty-one patients were assessed in this study (Table 2). There were 78 male (59.5%) and 53 female (40.5%) patients. The median age was 24. Seven patients experienced post-operative complications. Two patients were classed as Clavien grade II who developed a post-operative ileus; three patients classed as Clavien grade IIIa with post-operative collections and two patients classed as Clavien grade IIIb – one requiring a return to theatre for an inflammatory phlegmon and the other for a cecal volvulus. Of the two pregnant patients, one was in the first trimester and the other in the second trimester of pregnancy at the time of their presentation.

Table 2: Patient Demographics

Gender (F:M)	53:78
Age (years)	24 (16-31)ŧ
Length of stay (days)	2 (1-2)ŧ
Previous abdominal surgery	10¥
Post-operative complication	7¥
Pregnant	2¥
Negative appendectomy rate	11§

ŧ median (inter-quartile range IQR); ¥ absolute numbers; § percentage based on SAGS scoring

Table 3 demonstrates the agreement data for the surgical team. Linear-by-linear association test for ordinal data revealed a significant linear trend between pre-operative and intra-operative SAGS score (P = 0.001).

Weighted kappa (K_ω) for ordinal data (Table 4) showed K_ω index value of 0.62 for the surgical team, implying substantial association for pre-operative and intra-operative SAGS scores.

Table 3: Frequencies for ordered categories of severity of appendicitis: Surgical team SAGS[§]

Pre-operative SAGS score	Intra-operative SAGS score					
	0	**1**	**2**	**3**	**4**	**Total**
0	6	6	1	0	0	13
1	13	45	18	0	3	72
2	3	8	14	1	8	34
3	0	0	0	1	6	7
4	0	0	0	0	5	5
Total	22	59	26	2	22	131

↘ Diagonal agreement: perfect agreement between Pre- and Intra-operative SAGS § Linear trend between rows and columns P = 0.001 (Linear-by-linear association test)

Table 4: Pre-operative and intra-operative agreement index K_ω

Agreement:	Observed	Expected	K_ω	95% CI	p value
Surgical Team	94.27%	84.85%	0.62	0.51-0.74	0.001

K_ω interpretation: 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect

Log-linear modelling (Table 5) on the agreement data of Table 3 demonstrated that diagonal agreement plus linear-by-linear association to be the model of *best-fit* for the surgical team (deviance χ^2 7.56; degrees of freedom 14; P = 0.91).

Table 5: Log-linear modelling

Model	Deviance χ^2	*Degrees of freedom*	P value
Independence	96.457	16	< 0.001
Diagonal agreement	57.490	15	< 0.001
Quasi-independence	48.194	11	< 0.001
Linear-by-linear association	10.081	15	0.815
Diagonal agreement plus linear-by-linear association[¶]	7.564	14	0.911[¶]
Quasi-independence plus linear-by-linear association	6.149	10	0.803

Deviance χ^2: measure of lack-of-fit; null-hypothesis: the model is the correct model of best-fit relative to the baseline model of perfect fit; [¶] best-fit model.

Fig 1 demonstrates the relationship between intra-operative SAGS score with inflammatory markers. A Kruskal-Wallis test was used, suggesting a significant relationship between intra-operative SAGS score with CRP (P=0.0001), WCC (P=0.0001) and neutrophil count (P=0.0006).

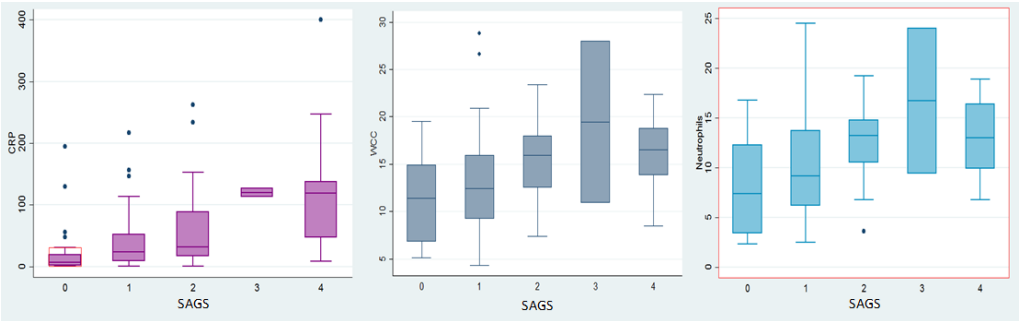

Figure 1: Relationship of SAGS score to CRP (mg/L), WCC (x10^9/L), and Neutrophil count (x10^9/L). Box-plots with median and inter-quartile range (IQR).

IV. Discussion

The SAGS score has previously been shown to be a clinically useful grading system to classify the intra-operative severity of appendicitis and guide antibiotic therapy. In this current study, pre-operative SAGS score in combination with routine laboratory tests have demonstrated a significant ability to assess intra-operative findings.

The commonly used kappa index measures the extent of observer agreement beyond that which would be expected by chance. The weighted kappa agreement index (K_ω) of 0.62 for the surgical team was interpreted as substantial, based on an arbitrary convention but provides no useful information about bias among raters [7-9]. In the agreement data of Table 3, ratings off the main diagonal are subject to systematic bias.[7-9]

The focus of the analysis in this study is the use of log-linear modelling for agreement data to examine the data for *lack-of-fit* between observed and expected frequencies and to find the model of *best-fit*. The *best-fit* model selected is indicated by a small deviance χ^2 value relative to the degrees of freedom [9].

The best-fit model of diagonal agreement plus linear-by-linear association (Table 5) indicated that SAGS scores were rated equally and similarly; high (or low) ratings made by one clinician tended to be associated with high (or low) ratings made by other clinicians. This is a more complete assessment of agreement (as distinct from association) than is possible from the use of weighted kappa.

The intra-operative SAGS score showed a significant trend with elevated CRP, WCC and neutrophilia (Fig. 1). There have been other studies to predict the severity of appendicitis using laboratory markers and imaging. However, no inflammatory marker alone such as WCC or CRP can predict the severity of appendicitis beset by varying levels of sensitivity and specificity [10,11]. A retrospective study of CT-imaging was reported to accurately diagnose the severity of appendicitis with reference to histological assessment, but this was at the cost of radiation exposure [12].

The negative appendectomy rate in this study was 12% based on pre-operative SAGS scoring. This was lower than the previously reported rate by Arthur et al., demonstrating a negative appendectomy rate of 19% across 27 Australian centers [13].

The strength of this study is its prospective design and its robust analysis of agreement data. However, there are some limitations. Firstly, it is a single-institution study. Secondly, it could be argued that scoring systems based on a structured data collection form has a checklist effect, a "Hawthorn effect" due to structured history taking and examination. This may be particularly relevant amongst trainees, resulting in a more consistent pre-operative patient assessment in terms of terminology and completeness of relevant items [14]. The authors acknowledge the use of an objective intra-operative grading system was applied to a subjective pre-operative assessment. However, this subjective assessment is reflective of the day-to-day practice of surgical clinicians asked to assess a patient with suspected appendicitis.

V. Conclusions

Pre-operative SAGS score assessment demonstrated a significant agreement with intra-operative findings, further supporting the clinical utility of the SAGS score. It has been shown to be a useful aid in the decision-making process in cases of suspected appendicitis and limit the use of pre-operative imaging. It remains a valuable tool to predict patients more likely to develop post-operative collections and to plan follow-up. This study also highlights that clinical assessment remains the most important tool to diagnosis appendicitis. The authors suggest the pre- and intra-operative use of the SAGS score to prioritize surgical resources and guide post-operative management.

Acknowledgements

Anthony T K Chan for advice and input on Log-linear modelling.

References

[1]. Jones PF, Suspected acute appendicitis: trends in management over 30 years, *British. Journal of Surgery, 88*, 2001, 1570-1577.
[2]. Shogilev DJ, Duus N, Odom SR, Shapiro NI, Diagnosing appendicitis: evidence-based review of the diagnostic approach in 2014. *Western Journal of Emergency Medicine, 15*, 2014, 859-871.
[3]. Alvarado A, A practical score for the early diagnosis of acute appendicitis, *Annuals of Emergency Medicine, 15*, 1986, 557-564.
[4]. Andersson M, Andersson RE, The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score, *World Journal of Surgery, 32*, 2008, 1843-1849.
[5]. Mikaere H, Zeng I, Lauti M, Kularatna M, MacCormick AD, Derivation and validation of the APPEND score: an acute appendicitis clinical prediction rule, *ANZ Journal of Surgery, 88*, 2018, E303-E307.
[6]. Reid F, Choi J, Williams M, Chan S, Prospective evaluation of the Sunshine Appendicitis Grading System score, *ANZ Journal of Surgery, 87*, 2017, 368-371.
[7]. Ludbrook J, Statistical techniques for comparing measurers and methods of measurement: a critical review. *Clinical and. Experimental Pharmacology and Physiology, 29*, 2002; 527–536.
[8]. Graham P, Jackson R, The analysis of ordinal agreement data: beyond weighted kappa, *Journal of Clinical Epidemiology, 46*, 1993, 1055-1062.
[9]. May SM, Modelling observer agreement – an alternative to kappa, *Journal of Clinical Epidemiology, 47*, 1994, 1315-1324.

[10]. Guraya SY, Al-Tuwaijri TA, Khairy, GA, Murshid, KR, Validity of leukocyte count to predict the severity of acute appendicitis, *Saudi Medical Journal*, *26*, 2005, 1945-1947.
[11]. van den Worm L, Georgiou E, de Klerk M, C-reactive protein as a predictor of severity of appendicitis, *South African Journal of Surgery*, *55*, 2017, 14-17.
[12]. Hansen AJ, Young SW, De Petris G, Tessier DJ, Hernandez JL, Johnson DJ, Histologic severity of appendicitis can be predicted by computed tomography, *The Archives of Surgery*, *139*, 2004, 1304–1308.
[13]. Arthur T, Gartrell R, Manoharan B, Parker D, QUEST Collaboration, Emergency appendicectomy in Australia: findings from a multicentre, prospective study, *ANZ Journal of Surgery*, *87*, 2017, 656–660
[14]. Korner H, Sondenaa K, Soreide JA, Andersen E, Nysted A, Lende TH, Structured data collection improves the diagnosis of acute appendicitis, *British Journal of Surgery*, *85*, 1998, 341-344.

Jacqueline Bakhit. "Pre-operative Assessment of Acute Appendicitis Severity with Reference to a Structured Intra-Operative Grading System." IOSR Journal of Dental and Medical Sciences (IOSR-JDMS), vol. 18, no. 2, 2019, pp 84-88.