9th August 2021

Dear Editor in Chief

World Journal of Critical Care Medicine

Re: Manuscript NO: 67039 **Sequential Organ Failure Assessment (SOFA) score is superior to other prognostic indices in the prediction of severity, intensive care unit admissions, and mortality in acute pancreatitis**

Many thanks for seeking peer reviews for our manuscript. We appreciate the time and expertise of peer reviewers and submit point by point response to the comments and have made edits in the manuscript.

Reviewer # 1

Comment 1: In the Abstract section, and throughout the manuscript (including in figure 1) please use "Ranson's" score rather than using "Ranson Score" interchangeably.

Response 1: Thank you for your comment. As suggested, we have elected to use "Ranson's" score as the term and have edited such throughout the paper.

Comment 2: In the Introduction section, I would recommend adding a one or two sentences how SOFA score is calculated. I suggest to provide specific references of the study that compared the efficacy of the newer studies rather than just quoting "Few studies…..have compared ….. Please write full form of APACHE in the introduction.

Response 2: Thank you for your suggestion. To address how SOFA score is calculated, we have added the following in the introduction:

The SOFA score is graded from 0 to 4 including markers PaO2/FiO2 ratio, Glasgow coma scale, mean arterial pressure or administration of vasopressors, bilirubin levels and platelet levels.

With respect to providing specific references to the studies comparing scoring systems, we have provided retrospective studies by Khanna et al and Tan et al comparing Ranson's score and Glasgow score, and Shafiq et al and Li et al. which provide critical clinical outcomes such as intensive care unit (ICU) admission and AP mortality. As below:

Khanna et al[14] and Tan et al[15] compared the efficacy of newer scores with traditional scores such as the Ranson's score and Glasgow score in predicting severity of AP.

Fewer still have reported their utility in predicting critical clinical outcomes such as intensive care unit (ICU) admission and AP mortality.

Further, the results are heterogenous as evidenced by the retrospective study by Shafiq et al[16] and Li et al[17]. A recent review critically appraised the current evidence regarding clinical utility of various scoring systems in AP, and concluded that "old is gold" in favor of traditional scoring system like Ranson's score[18].

We have also inserted the full form of "APACHE-II" in the introduction.

Comment 3: In Methods section:

a. Could you please provide the reason of taking only the patients admitted in the surgery department? Were Alcoholic pancreatitis admitted preferably in medicine department? Alcoholic pancreatitis generally cause severe pancreatitis- so they might be excluded from the study. Is there possibility we missed significant cases of Alcoholic AP?

b. SOFA score calculation: Was SOFA score calculated once only? If repeated after 48 hours, which SOFA score (mean or the highest) value was used?

Response 3: Thank you for your comments.

Regarding alcoholic pancreatitis

In this study, incidence of alcoholic pancreatitis was 5.8%. This is likely due to two reasons: (a) low prevalence of alcoholic pancreatitis in local population, and (b) some alcoholic pancreatitis patients with coincidental gallstones and clinically not possible to distinguish exact aetiology,

were assigned 'gallstone' aetiology. This is done to ensure that counseling for laparoscopic cholecystectomy is not missed out at follow up visits.

The local emergency department admits all patients with acute abdominal pain to general surgery department. Thus, it is unlikely that we selectively missed alcoholic pancreatitis patients. It is possible that a minority of patients who were admitted to general medicine department for other symptoms might be diagnosed with acute pancreatitis after a few days of admission, and these patients could be missed. However, locally, once a patient is diagnosed with acute pancreatitis, a surgical referral is made and many times, the patient is transferred back to surgical department. Thus, the low incidence (5.8%) of alcoholic pancreatitis is unlikely due to missing out selectively patients with alcohol as aetiology. However, as the possibility exists that patients admitted to general medicine department, and with diagnosis of acute pancreatitis could be missed in this data, we have acknowledged this as a limitation of our study.

Patients admitted to general medicine department were excluded in this study, and thus it is possible that true prevalence of alcoholic pancreatitis is under reported.

Regarding SOFA score

This is a very important comment. It is proven that dynamic score and trends are superior predictor of severity of illness e.g. worsening trends are more worrisome in clinical context than single worst index. This study is a retrospective medical record review, and it is in essence an audit of local data. We have computed the SOFA and all other scores at single time point, including the patient physiology and laboratory values within the first 24 hours of admission. The time of admission is distinct from time of onset of acute pancreatitis, which is determined by onset of abdominal pain. Thus, it is not possible to compute a score based on time of onset of acute pancreatitis, which probably is a more accurate reflection of performance of a scoring system. In the method section, we have modified the statement to include the 'worst recorded data point'. It now reads as:

Values for APACHE II score, BISAP, HAPS, and SOFA within 24 hours of admission were retrospectively obtained based on the worst recorded laboratory results and patient vital charts.

Timing of score computation is very integral to performance, and thus we have discussed to acknowledge this as a limitation of our study.

In addition, onset of abdominal pain is different than the date of hospitalization. Thus, computing a score based on parameters within 24 hours of admission does not necessarily provide an accurate reflection of score performance, as acute pancreatitis management timelines is determined by onset of abdominal pain and not time of hospitalization.

To ensure consistency in reporting standards, department has a policy to compute Ranson's score and Glasgow score at 48 hours.

Comment 4: In the Definitions, I would similarly recommend substituting "By current guidelines" with According to Revised Atlanta guidelines, in the Study Outcomes section.

In prognostics scoring section of Definitions, I suggest either to add year in all scorings or drop the year. The development is mentioned for Ranson's, Glasgow score but not for APACHE, BISAP, HAPS, SOFA.

Response 4: Thank you for your suggestion. We have edited the Study Outcomes to include Revised Atlanta guidelines as follows: *According to the revised Atlanta guidelines, AP can be graded as mild, moderately severe, or severe.*

We have also taken your advice and removed all the years of conception of scoring systems for consistency:

Ranson's score was the first developed ~~in 1974~~ to risk-stratify AP[8] and consists of 11 parameters, five scored at admission, and six scored at 48 hours after admission. Glasgow score, otherwise known as the Glasgow-Imrie or Imrie score, was first described by Blamey *et al.* ~~in 1984~~[16] and consist of eight variables scored with values at 48 hours after admission.

Comment 5: In the statistical analysis section, third line - Continuous variables… it says mean ± standard deviation represents an interquartile range (IQR), but interquartile range in difference between first and third quartile. Could you please make it clear?

Response 5: Thank you for your observation. We acknowledge the error. We have revised the statement to remove the "IQR".

Categorical variables are presented as absolute numbers and proportions. Continuous variables are presented as mean ± standard deviation (SD). representing an interquartile range

Comment 6: The discussion section seems denser, so I suggest it to make it more coherent by breaking into subheadings and removing the information that are not particularly important i.e, CI when mentioning AUROC value.

Response 6: Thank you for your comment. We agree that providing subheadings would provide more structure to the paper and have added subheadings consistent with our "Results" section for convenience including:

Patient characteristics
Prediction of Severe Acute Pancreatitis (SAP)
Intensive Care Unit (ICU) Admission
Mortality in Acute Pancreatitis
Limitations

Information that was deemed not particularly important was removed as well as suggested, consistent with Comment 8 as well.

We have also removed the CI when mentioning the AUROC values in discussion section.

Comment 7: In the second paragraph of discussion section, last sentence- The lower prevalence of alcoholic pancreatitis may reflect lower consumption rates in the Asian population. Could this be also due to Alcoholic pancreatitis admitted in other departments and those were not included in the study?

Response 7: Thanks for this comment. In this study, incidence of alcoholic pancreatitis was 5.8%. This is likely due to two reasons: (a) low prevalence of alcoholic pancreatitis in local population, and (b) some alcoholic pancreatitis patients with coincidental gallstones and clinically not possible to distinguish exact aetiology, were assigned 'gallstone' aetiology. This is done to ensure that counseling for laparoscopic cholecystectomy is not missed out at follow up

visits. The local emergency department admits all patients with acute abdominal pain to general surgery department. Thus, it is unlikely that we selectively missed alcoholic pancreatitis patients.

We have added the following in discussion section: *Some alcoholic AP patients with coincidental gallstones might be assigned 'gallstone' aetiology, so that laparoscopic cholecystectomy recommendation is not missed out at follow up visits.*

It is possible that a minority of patients who were admitted to general medicine department for other symptoms might be diagnosed with acute pancreatitis after a few days of admission, and these patients could be missed. However, locally, once a patient is diagnosed with acute pancreatitis, a surgical referral is made and many times, the patient is transferred back to surgical department. Thus, the low incidence (5.8%) of alcoholic pancreatitis is unlikely due to missing out selectively patients with alcohol as aetiology. However, as the possibility exists that patients admitted to general medicine department, and with diagnosis of acute pancreatitis could be missed in this data, we have acknowledged this as a limitation of our study.

Patients admitted to general medicine department were excluded in this study, and thus it is possible that true prevalence of alcoholic pancreatitis is under reported.

Comment 8: In discussion section please drop the confidence interval while mentioning the point estimate of the score value.

Response 8: Consistent with Comment 6, we have removed the confidence interval ranges for the point estimate of the score values.

Comment 9: In Figures and legends:

Figure 1,2,3: Please clarify what is Ranson's Cumulative? Is it Ransons's at 48 hours? Please use same terminology in ROC figure and table. It is better to use only figure rather than combing with table and labelling it as figure.

Response 9: Thank you for your comment. We have replaced Ranson's culmulative with Ranson's score at 48 hours for consistency. We have removed the table in favor of the figure as suggested.

Comment 10: Table 1: Could you please mention what was included in the others and idiopathic etiology?

Response 10: Thank you for your comment. We have added below Table 1 the elaboration:
For idiopathic etiology, we have defined it as Acute Pancreatitis with no etiology despite extensive work up:
Idiopathic = Acute Pancreatitis with no etiology despite extensive work up
For 'others': *Others = Etiologies of acute pancreatitis include trauma, pancreas cystic neoplasms, malignancy, iatrogenic causes such as Endoscopic retrograde cholangiopancreatography (ERCP).*

In addition, we realised that we should better include aetiology section in study methodology. This we have added a section on diverse aetiology of acute pancreatitis, including definitions. The following is added:

Aetiology of Acute Pancreatitis

Gallstones, alcohol, hypertriglyceridemia, hypercalcemia, trauma, pancreas cystic neoplasms, post endoscopic retrograde cholangiopancreatography (ERCP), autoimmune, and drug-induced pancreatitis were common identified aetiologies. Gallstone-associated AP was defined when imaging detected gallstones. Alcohol-associated AP was based on the presence of a recent alcoholic binge or regular high intake, and absence of gallstones. Hypertriglyceridemia-associated AP was diagnosed if serum TG was >10 mmol/L. Hypercalcemia-associated AP was diagnosed if serum ionized calcium was >1.3 mmol/L. Drug-induced AP was diagnosed in the absence of common aetiologies and in the presence of recent intake of known culprit medications like hydrochlorothiazide, statins, non-steroidal anti-inflammatory drugs (NSAIDs), etc. Idiopathic AP was diagnosed when extensive diagnostic work-up including magnetic resonance cholangiopancreatography (MRCP) scan, endoscopic ultrasound (EUS), and serum immunoglobulin tests failed to establish a clear aetiology.

Comments from Science Editor:

Comment 1. The authors did not provide original pictures. Please provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure that all graphs or arrows or text portions can be reprocessed by the editor;

Response 1: We have provided the necessary figures in PowerPoint slides as attached.

Comment 2. The title is too long, and it should be no more than 18 words

Response 2: We have shortened the title to: **Sequential Organ Failure Assessment (SOFA) score is superior to other prognostic indices in acute pancreatitis**

Comment 3. (3) The "Article Highlights" section is missing. Please add the "Article Highlights" section at the end of the main text; And

Response 3: We have added the following under "Article Highlights":

Article Highlights

(1) Research background

Acute Pancreatitis (AP) is a common surgical condition, with severe AP (SAP) potentially lethal. Many prognostic indices, including; Acute physiology and chronic health evaluation II score (APACHE II), Bedside index of severity in acute pancreatitis (BISAP), Glasgow score, Harmless acute pancreatitis score (HAPS), Ranson's score, and Sequential organ failure assessment (SOFA) evaluate AP severity and predict mortality.

(2) Research motivation

An accurate scoring system on admission is critical to guide AP patient disposition and aggressiveness of treatment, resulting in both better patient care as well as better distribution of healthcare resources. Few studies have compared the efficacy of these newer scores in predicting disease severity against classic scores such as Ranson's score and Glasgow score, and fewer still have reported their utility in predicting key clinical outcomes such as intensive care unit (ICU) admission and mortality in AP.

(3) Research objectives

A major concern for clinicians is the gross heterogeneity in clinical presentation and identifying patients predicted to manifest severe AP (SAP). We evaluated these indices' utility in predicting severity, Intensive Care Unit (ICU) admission, and mortality.

(4) Research methods

A retrospective analysis of 653 patients with AP from July 2009 to September 2016 was performed. The demographic, clinical profile, and patient outcomes were collected. SAP was defined as per the revised Atlanta classification. Values for APACHE II, BISAP, HAPS, and SOFA score within 24 hours of admission were retrospectively obtained. Variables with <10% missing data were imputed via mean substitution.

(5) Research results

RESULTS

The mean age was 58.7±17.5 years, with 58.7% males. Gallstones (n=404, 61.9%), alcohol (n=38, 5.8%), and hypertriglyceridemia (n=19, 2.9%) were more common aetiologies. 81(12.4%) patients developed SAP, 20(3.1%) required ICU admission, and 12(1.8%) deaths were attributed to SAP. Ranson's score and APACHE-II demonstrated the highest sensitivity in predicting SAP (92.6%, 80.2% respectively), ICU admission (100%), and mortality (100%). SOFA and BISAP scores demonstrated lowest sensitivity in predicting SAP (13.6% and 24.7% respectively), ICU admission (40.0% and 25.0% respectively) and mortality (50.0% and 25.5% respectively). However, SOFA demonstrated the highest specificity in predicting SAP (99.7%), ICU admission (99.2%), and mortality (98.9%). SOFA demonstrated the highest positive predictive value, positive likelihood ratio, diagnostic odds ratio, and overall accuracy in predicting SAP, ICU admission, and mortality. SOFA and Ranson's score demonstrated the highest Area under receiver-operator curves (AUROC) at 48 hours in predicting SAP (0.966 and 0.857 respectively), ICU admission (0.943 and 0.946 respectively), and mortality (0.968 and 0.917 respectively).

(6) Research conclusions

Overall, the six prognostic indices in this study demonstrated high negative predictive values in prediction of severity, ICU admission and mortality in AP. SOFA score and Ranson's score at 48 hours

are superior to other prognostic scorings (Glasgow score, APACHE II, BISAP, HAPS) in severity stratification, prediction of ICU admission and mortality in AP.

(7) Research perspectives

As we provide a retrospective single-center study, future renditions of this study could include international collaborative multi-center analysis to increase validity. Each healthcare unit should perform regular audits to identify most accurate and relevant scoring system for local population. No one scoring system is perfect. Further studies may compare the utility of trends of different scores throughout inpatient stay.

Comment 4: Please re-write the "Abstract" section according to the Guidelines and Requirements for Manuscript Revision.

Response 4: As requested, we have re-written the "Abstract" section in accordance to the Guidelines and Requirements for Manuscript Revision:

ABSTRACT

BACKGROUND

Acute Pancreatitis (AP) is a common surgical condition, with severe AP (SAP) potentially lethal. Many prognostic indices, including; Acute Physiology And Chronic Health Evaluation II score (APACHE II), Bedside Index of Severity in Acute Pancreatitis (BISAP), Glasgow score, Harmless Acute Pancreatitis Score (HAPS), Ranson's score, and Sequential Organ Failure Assessment (SOFA) evaluate AP severity and predict mortality.

AIM

We evaluated these indices' utility in predicting severity, Intensive Care Unit (ICU) admission, and mortality.

METHODS

A retrospective analysis of 653 patients with AP from July 2009 to September 2016 was performed. The demographic, clinical profile, and patient outcomes were collected. SAP was defined as per the revised Atlanta classification. Values for APACHE II score, BISAP, HAPS, and SOFA within 24 hours of admission were retrospectively obtained. Variables with <10% missing data were imputed via mean substitution. Other patient information such as demographics, disease etiology, and patient outcomes were also derived from electronic medical records.

RESULTS

The mean age was 58.7±17.5 years, with 58.7% males. Gallstones (n=404, 61.9%), alcohol (n=38, 5.8%), and hypertriglyceridemia (n=19, 2.9%) were more common aetiologies. 81(12.4%) patients developed SAP, 20(3.1%) required ICU admission, and 12(1.8%) deaths were attributed to SAP. Ranson's score and APACHE-II demonstrated the highest sensitivity in predicting SAP (92.6%, 80.2% respectively), ICU admission (100%), and mortality (100%). While SOFA and BISAP demonstrated lowest sensitivity in predicting SAP (13.6%, 24.7% respectively), ICU admission (40.0%, 25.0% respectively) and mortality (50.0%, 25.5% respectively). However, SOFA demonstrated the highest specificity in predicting SAP (99.7%), ICU admission (99.2%), and mortality (98.9%). SOFA demonstrated the highest positive predictive value, positive likelihood ratio, diagnostic odds ratio, and overall accuracy in predicting SAP, ICU admission, and mortality. SOFA and Ranson's score demonstrated the highest Area under Receiver-operator Curves (AUROC) at 48 hours in predicting SAP (0.966, 0.857 respectively), ICU admission (0.943, 0.946 respectively), and mortality (0.968, 0.917 respectively).

CONCLUSION

The SOFA and 48-hour Ranson's scores accurately predict severity, ICU admission, and mortality in AP, with more favorable statistics for the SOFA score.

Thank you

Sincerely

Vishal G Shelat,

Associate Professor

Department of General Surgery

Tan Tock Seng Hospital 308433