PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association between neutrophil-to-albumin ratio and mortality in patients with cardiogenic shock: a retrospective cohort study
AUTHORS	Peng, Yangpei; Xue, Yangjing; wang, Jinsheng; Xiang, Huaqiang; ji, kangting; Wang, Jie; Lin, Cong

VERSION 1 – REVIEW

REVIEWER	Sebastian Reinstadler Medical University of Innsbruck, Austria
REVIEW RETURNED	05-May-2020

GENERAL COMMENTS	Yangpei et al performed a retrospective cohort analysis (n=475) evaluating the association between neutrophil-to-albumin ratio (NAR) and mortality in patients with cardiogenic shock. The primary clinical outcome was 90-day mortality. The authors observed higher rates of 90-day mortality (as well as 30-day and 365-day mortality) in patients with high NAR values (according to NAR tertiles). The predictive value of NAR was moderate (AUC 0.65). In multivariable analysis, NAR remained associated with mortality. I have the following specific comments: Major concern is on whether NAR is incremental to that of established parameters (SOFA, lactate, etc.). This point should be elaborated much further. First, the authors should compare AUCs and test for statistical significant differences between AUCs. Second, while a classification of NAR by tertiles seems appropriate, an additional analysis showing the significant and independent relationship with the primary outcome using NAR as a continuous variable would be important. Moreover, what happens after categorizing according to NAR median? Third, previously described parameters of poor outcome (mainly lactate) in patients with cardiogenic shock (Pöss et al JACC 2017) should be considered. The authors should more clearly state at which time point NAR was determined. Minor - Abstract: Write out abbreviations the first time - Abstract: Please specify at which time point NAR was determined - Following Rice and Harris, AUC values were categorized as negligible (≤0.55), small (0.56–0.63), moderate (0.64–0.70), and strong (≥0.71). The authors should follow this classification

REVIEWER	Kim Oren Gradel

	Center for Clinical Epidemiology Odense University Hospital Denmark
REVIEW RETURNED	14-May-2020

GENERAL COMMENTS	This prognostic cohort study assesses 30-day, 90-day, and 365-day mortality after cardiogenic shock, where the main exposure is the neutrophil-to-albumin ratio. For some reason, 90-day is chosen as the primary outcome, but I need a justification for that. As the analyses are made, 90-day does not cover long-term outcome, it actually covers both short-term (0-30 days) and long-term (31-90 days) outcome. A more straightforward way would be to divide the periods into 0-30, 31-90, and 91-365 days, which would really reflect the different periods. As it is now, it is probably the short-term mortality (0-30 days) which mainly contributes to the significance for the 90- and 365-day analyses. This is also indicated in table 1 where no. deaths are decreasing considerably after day 30. To quote: “A total of 180, 218 and 264 deaths were recorded in the 30-day, 90-day and 365-day follow-up periods, respectively”. This gives a mean of 6 deaths per day on day 0-30, 0.6 per day on day 31-90, and 0.17 per day on day 91-365. Kaplan-Meier curves would be good to show this. Moreover, it is indicated in figures 2 and 3, where the curves are very similar for 30-day, 90-day, and 365-day mortality. There is an overwhelming number of subgroup analyses. Maybe these would suit better in a supplementary file. Fig. 4 and 5 indicate that NAR contributes very little to the already known and well-consolidated prognostic indices SAPSII and SOFA and the clinical relevance of the low AUCs may be discussed. So I disagree with the interpretation that “Building further on this foundation, combining NAR and SOFA score or SAPS II can provide greater prognostic values than NAR alone, just as Figure 4 and 5 showed.” NAR is divided into both tertiles and quintiles, what is the purpose of that, it confuses the reader? If there is a clear trend, why bother to categorize this continuous variable? I do not agree that dividing into quintiles is a strength of the study. I wonder about the low AUC of albumin (0.509), which is like tossing a coin. This is in contrast to numerous other studies that indicate that albumin per se is a strong prognostic predictor in numerous patient groups [1, 2]. What is the justification for the exclusion criteria (p. 6)? Why not just include all and then perform subgroup analyses where these are excluded to assess the robustness and generality of the results? A lot of biomarkers are included, but few biomarkers that directly assess inflammation, e.g. CRP or PCT, which is strange given that inflammation is described as a key point in CS. I do not understand the sentence “Follow-up began when the patients first admitted to the ICU and stayed for at least one year”. Did they stay for 1 y in the ICU?

	I do not understand the sentence “their associations with mortality or a mutation exceeding 10%”. Model II uses a lot of variables, probably too many given possible collinearities and the numbers of outcomes (rule of thumb: one independent variable per 10 outcomes). I do not understand the sentence “As blood cells and serum albumin can directly indicate the host environment” Quote: “It has already become a part of major risk scores, such as the Acute Physiology and Chronic Health Evaluation (APACHE) III Prognostic system and the Framingham risk score for cardiovascular risk.” What are the references for that? I guess that it is [3] for APACHE III, but I do not think that albumin is part of the Framingham risk score. If I am wrong, please give a reference. If I am correct, please delete this. Quote: “Another interpretation may involve the state of nutrition because lower albumin is possibly related to malnutrition, emaciation or cachexia”: There are so many reviews and studies that indicate that albumin is a lousy nutrition marker, e.g. [1, 4-6], so I tend to disagree with this statement. In the tables, why not trend tests instead of tests that “just” show differences between two or more groups? References [1] Levitt DG, Levitt MD. Human serum albumin homeostasis: a new look at the roles of synthesis, catabolism, renal and gastrointestinal excretion, and the clinical value of serum albumin measurements. Int J Gen Med 9;2016:229-55. [2] Vincent JL, Dubois MJ, Navickis RJ, Wilkes MM. Hypoalbuminemia in acute illness: is there a rationale for intervention? A meta-analysis of cohort studies and controlled trials. Ann.Surg. 237;2003:319-334. [3] Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 100;1991:1619-1636. [4] Lee JL, Oh ES, Lee RW, Finucane TE. Serum albumin and prealbumin in calorically restricted, nondiseased individuals: a systematic review. Am J Med 128;2015:1023 e1-22. [5] Fuhrman MP. The albumin-nutrition connection: separating myth from fact. Nutrition 18;2002:199-200. [6] Soeters PB, Wolfe RR, Shenkin A. Hypoalbuminemia: Pathogenesis and Clinical Significance. JPEN J Parenter Enteral Nutr 43;2019:181-193.

VERSION 1 – AUTHOR RESPONSE

Responses to Reviewer: 1
Yangpei et al performed a retrospective cohort analysis (n=475) evaluating the association between neutrophil-to-albumin ratio (NAR) and mortality in patients with cardiogenic shock. The primary clinical outcome was 90-day mortality. The authors observed higher rates of 90-day mortality (as well as 30-day and 365-day mortality) in patients with high NAR values (according to NAR tertiles). The predictive value of NAR was moderate (AUC 0.65). In multivariable analysis, NAR remained associated with mortality.
Thank you for your comments.

I have the following specific comments:
Major concern is on whether NAR is incremental to that of established parameters (SOFA, lactate, etc.). This point should be elaborated much further. First, the authors should compare AUCs and test for statistical significant differences between AUCs. Second, while a classification of NAR by tertiles seems appropriate, an additional analysis showing the significant and independent relationship with the primary outcome using NAR as a continuous variable would be important. Moreover, what happens after categorizing according to NAR median? Third, previously described parameters of poor outcome (mainly lactate) in patients with cardiogenic shock (Pöss et al JACC 2017) should be considered.
Thank you for your comments. First, we have compared and tested the statistically significant differences of AUCs, and the differences were statistically significant. Second, the analysis showing the significant and independent relationship with the outcomes using NAR as a continuous variable has been added. Third, the data extracted from the MIMIC database were all shown in our study, except for those with too many missing values that cannot be analyzed, including lactate.

The authors should more clearly state at which time point NAR was determined.
Thank you for your comments. We have clearly stated the time at which NAR was determined.

Minor
- Abstract: Write out abbreviations the first time
Thank you for your comments. We have revised them.

- Abstract: Please specify at which time point NAR was determined
Thank you for your comments. We have added the time at which NAR was determined.

- Following Rice and Harris, AUC values were categorized as negligible (≤0.55), small (0.56–0.63), moderate (0.64–0.70), and strong (≥0.71). The authors should follow this classification
Thank you for your comments. We have followed the classification of AUC values.

Responses to Reviewer: 2
This prognostic cohort study assesses 30-day, 90-day, and 365-day mortality after cardiogenic shock, where the main exposure is the neutrophil-to-albumin ratio.
Thank you for your comments.

For some reason, 90-day is chosen as the primary outcome, but I need a justification for that.
Thank you for your comments. We chose 90-day mortality as the primary outcome according to the methods of relevant published studies.

As the analyses are made, 90-day does not cover long-term outcome, it actually covers both short-term (0-30 days) and long-term (31-90 days) outcome. A more straightforward way would be to divide the periods into 0-30, 31-90, and 91-365 days, which would really reflect the different periods. As it is now, it is probably the short-term mortality (0-30 days) which mainly contributes to the significance for

the 90- and 365-day analyses. This is also indicated in table 1 where no. deaths are decreasing considerably after day 30. To quote: “A total of 180, 218 and 264 deaths were recorded in the 30-day, 90-day and 365-day follow-up periods, respectively”. This gives a mean of 6 deaths per day on day 0-30, 0.6 per day on day 31-90, and 0.17 per day on day 91-365. Kaplan-Meier curves would be good to show this. Moreover, it is indicated in figures 2 and 3, where the curves are very similar for 30-day, 90-day, and 365-day mortality.
Thank you for your constructive comments. We have carefully reviewed our results and revised our manuscript.

There is an overwhelming number of subgroup analyses. Maybe these would suit better in a supplementary file.
Thank you for your comments. We have resubmitted the most results of subgroup analyses as a supplementary file.

Fig. 4 and 5 indicate that NAR contributes very little to the already known and well-consolidated prognostic indices SAPSII and SOFA and the clinical relevance of the low AUCs may be discussed. So I disagree with the interpretation that “Building further on this foundation, combining NAR and SOFA score or SAPS II can provide greater prognostic values than NAR alone, just as Figure 4 and 5 showed.”
Thank you for your comments. We are very sorry for the mistake in our expression. We have revised it.

NAR is divided into both tertiles and quintiles, what is the purpose of that, it confuses the reader? If there is a clear trend, why bother to categorize this continuous variable? I do not agree that dividing into quintiles is a strength of the study.
Thank you for your comments. We have revised it.

I wonder about the low AUC of albumin (0.509), which is like tossing a coin. This is in contrast to numerous other studies that indicate that albumin per se is a strong prognostic predictor in numerous patient groups [1, 2].
Thank you for your comments. We are sorry that we have reversed the AUC values for neutrophil and albumin in the manuscript. The AUC of albumin was 0.584, and that of neutrophil was 0.509, just as shown in Figure 2. We have revised them.

What is the justification for the exclusion criteria (p. 6)? Why not just include all and then perform subgroup analyses where these are excluded to assess the robustness and generality of the results?
Thank you for your comments. The exclusion criteria were identified referring to the methods of relevant published studies.

A lot of biomarkers are included, but few biomarkers that directly assess inflammation, e.g. CRP or PCT, which is strange given that inflammation is described as a key point in CS.
Thank you for your comments. The data in our study were all extracted from the MIMIC database. Because some biomarkers (e.g. CRP) have too many missing values, we fail to analyze them further. As for PCT, there was no relevant data in the database.

I do not understand the sentence “Follow-up began when the patients first admitted to the ICU and stayed for at least one year”. Did they stay for 1 y in the ICU?
Thank you for your comments and we are sorry for the misunderstanding caused by the typo. We have revised it.

I do not understand the sentence “their associations with mortality or a mutation exceeding 10%”.
Thank you for your comments. This was the method we applied to select the confounders. In

multivariate model, the confounders were selected based on their associations with the outcome or they caused more than 10% of the mutations in the model. This method also borrowed from previously published articles.

Model II uses a lot of variables, probably too many given possible collinearities and the numbers of outcomes (rule of thumb: one independent variable per 10 outcomes).
Thank you for your comments. The statistical analysis method of our study was based on previous similar studies. The adjusted variables used in the multivariate model were based on their associations with the outcome or they caused more than 10% of the mutations in the model.

I do not understand the sentence “As blood cells and serum albumin can directly indicate the host environment”
Thank you for your comments. We are sorry for our sloppy presentation. We have revised it.

Quote: “It has already become a part of major risk scores, such as the Acute Physiology and Chronic Health Evaluation (APACHE) III Prognostic system and the Framingham risk score for cardiovascular risk.” What are the references for that? I guess that it is [3] for APACHE III, but I do not think that albumin is part of the Framingham risk score. If I am wrong, please give a reference. If I am correct, please delete this.
Thank you for your comments. We have deleted it.

Quote: “Another interpretation may involve the state of nutrition because lower albumin is possibly related to malnutrition, emaciation or cachexia”: There are so many reviews and studies that indicate that albumin is a lousy nutrition marker, e.g. [1, 4-6], so I tend to disagree with this statement.
Thank you for your comments. We are sorry for our sloppy presentation. We have revised it.

In the tables, why not trend tests instead of tests that “just” show differences between two or more groups?
Thank you for your comments. We have added the P values of trend tests.

References
[1] Levitt DG, Levitt MD. Human serum albumin homeostasis: a new look at the roles of synthesis, catabolism, renal and gastrointestinal excretion, and the clinical value of serum albumin measurements. Int J Gen Med 9;2016:229-55.
[2] Vincent JL, Dubois MJ, Navickis RJ, Wilkes MM. Hypoalbuminemia in acute illness: is there a rationale for intervention? A meta-analysis of cohort studies and controlled trials. Ann.Surg. 237;2003:319-334.
[3] Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 100;1991:1619-1636.
[4] Lee JL, Oh ES, Lee RW, Finucane TE. Serum albumin and prealbumin in calorically restricted, nondiseased individuals: a systematic review. Am J Med 128;2015:1023 e1-22.
[5] Fuhrman MP. The albumin-nutrition connection: separating myth from fact. Nutrition 18;2002:199-200.
[6] Soeters PB, Wolfe RR, Shenkin A. Hypoalbuminemia: Pathogenesis and Clinical Significance. JPEN J Parenter Enteral Nutr 43;2019:181-193.
Thank you very much for providing the above references, which are very helpful to us.

VERSION 2 – REVIEW

REVIEWER	Sebastian Reinstadler

	Medical University of Innsbruck Austria
REVIEW RETURNED	09-Jul-2020

GENERAL COMMENTS	Many thanks for the revision of the manuscript. Many parts have been adequately revised. However, the main concern remains as to whether NAR provides any additional information on the well-known risk predictors. In fact, the combination of NAR and SOFA or SAPSII does not result in a significant increase in AUC (Figure 3 and 4). In addition, when NAR was modeled as a continuous variable it seems that it is no longer independently associated with the primary endpoint (see Model 2 in Table 2). The authors should provide clear evidence for supporting their conclusions that “NAR was an independent predictor of mortality in CS patients. As a readily-available biomarker, NAR may be potential for risk stratification of CS.” Otherwise, these conclusions are not fully supported by the data and should therefore be revised accordingly.

REVIEWER	Kim Oren Gradel Center for Clinical Epidemiology, Odense University Hospital, Denmark
REVIEW RETURNED	23-Jul-2020

GENERAL COMMENTS	The authors have generally adhered to my recommendations, of which the most important was the little contribution of NAR to the area under the ROC curve. The following comments refer to some of my concerns that have not been addressed, but most of these are minor: Regarding 0-30, 0-90, and 0-365 mortality (instead of 0-30, 31-90, and 91-365 day mortality): The authors write they have revised this, but that does not seem to be the case. Still, most articles on different survival time spans use these methods, so I am a minority here. I still do not understand the word “mutation”. You refer to ref. 23, but that article does not include the word mutation. It is a genetic term, I have never encountered it in statistics. After having seen ref. 23 I understand what you mean, but please you the proper term. My suggestions for trend tests also relates to Table 1 and Table 3. Moreover, the trend test has not been described in the Methods section although they appear in table 2. In table 3 you mention p for interaction. What is that? Interaction is a synonym for effect modification, and this is not what you are testing for…

VERSION 2 – AUTHOR RESPONSE

Responses to Reviewer: 1

Many thanks for the revision of the manuscript. Many parts have been adequately revised. However, the main concern remains as to whether NAR provides any additional information on the well-known risk predictors. In fact, the combination of NAR and SOFA or SAPSII does not result in a significant increase in AUC (Figure 3 and 4). In addition, when NAR was modeled as a continuous variable it seems that it is no longer independently associated with the primary endpoint (see Model 2 in Table 2). The authors should provide clear evidence for supporting their conclusions that “NAR was an independent predictor of mortality in CS patients. As a readily-available biomarker, NAR may be

potential for risk stratification of CS." Otherwise, these conclusions are not fully supported by the data and should therefore be revised accordingly.
Thank you for your comments. Although the combination of NAR and SOFA or SAPSII does not result in a significant increase in AUC, NAR is more readily available than SOFA or SAPSII and can play a role in rapid clinical assessment. In addition, our results showed that NAR is more sensitive in predicting mortality of CS than the neutrophil percentage or the serum albumin level alone. For the model using NAR as a continuous variable, we are sorry that the results of 90-day and 30-day were reversed. NAR was still independently associated with the primary endpoint in Model II. As for the conclusions, we are lack of rigour and we have revised them.

Responses to Reviewer: 2
The authors have generally adhered to my recommendations, of which the most important was the little contribution of NAR to the area under the ROC curve. The following comments refer to some of my concerns that have not been addressed, but most of these are minor:
Regarding 0-30, 0-90, and 0-365 mortality (instead of 0-30, 31-90, and 91-365 day mortality): The authors write they have revised this, but that does not seem to be the case. Still, most articles on different survival time spans use these methods, so I am a minority here.
Thank you for your comments. Actually, the aim of our study was just to investigate the association between NAR and mortality of CS patients. The outcomes were selected referring to previous similar studies. We agree that the prognostic value of NAR in the different periods of CS can be more directly investigated by dividing the periods into 0-30, 31-90, and 91-365 days. However, our sample size of included population was relatively small. We believe a deeper study of a larger sample would make sense.

I still do not understand the word "mutation". You refer to ref. 23, but that article does not include the word mutation. It is a genetic term, I have never encountered it in statistics. After having seen ref. 23 I understand what you mean, but please you the proper term.
Thank you for your comments. We have revised it.

My suggestions for trend tests also relates to Table 1 and Table 3. Moreover, the trend test has not been described in the Methods section although they appear in table 2.
Thank you for your comments. The statistical analyses in our study were based on previous similar studies. For Table 1, we did comparisons between groups by the chi-square test or Fisher's exact test for categorical variables and the variance analysis or the Kruskal-Wallis test for continuous ones. For Table 3, we just compared the difference between the subgroups.
Previous similar studies:
DOI: 10.1186/s13054-017-1821-z
DOI: 10.1016/j.cca.2018.09.014
In addition, we have described the trend test in the Methods section.

In table 3 you mention p for interaction. What is that? Interaction is a synonym for effect modification, and this is not what you are testing for…
Thank you for your comments. We have revised it.

VERSION 3 – REVIEW

REVIEWER	Sebastian Reinstadler Medical University of Innsbruck
REVIEW RETURNED	24-Aug-2020

GENERAL COMMENTS	Thanks for the thorough revisions. I have no further comments.