PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and validation of a prognostic nomogram for myocardial infarction patients in intensive care units: a retrospective cohort study
AUTHORS	Guo, Qi; Wu, Maoxiong; Li, Hongwei; Ouyang, Huijun; Sun, Runlu; Wang, Junjie; Liu, Zhaoyu; Wang, Jingfeng; Zhang, Yuling

VERSION 1 – REVIEW

REVIEWER	Carmelo Dominici Università Campus Bio-Medico di Roma, Rome, Italy
REVIEW RETURNED	07-Jul-2020

GENERAL COMMENTS	authors should be congratulated for using a time-to-event analysis. however, considering that "30-days mortality" (a binary variable) was the primary endpoint, the rationale of choosing a Cox regression analysis should be described. proportional hazard assumption should be checked for each variable included in the (univariable) model when using Cox analysis. alternatively, logistic regression can be more easily performed to detect significant predictive variables. please discuss the rationale of the cut-off for inclusion in multivariable analysis. generally a P value of 0.1-0.2 in univariable analysis is used. from the manuscript, it appears that the cut-off was 0.05. survival curves should include 95% confidence intervals, number at risk for each group and censored cases. please include those 2 figures as separate images. in results, please avoid text/table duplication. include most numeric results in tables to improve fluency of the manuscript. please provide 2 figures with examples of a low risk patient and a high risk patient (e.g. with numbers provided in page 9, lines 38-46). differences in threshold probabilities between derivation cohort and validation cohort should be discussed. a table comparing this nomogram with existing models (e.g. number of patients included in cohorts, AUC, predictive probabilities,...) should be provided. some factors of clinical presentation are known to affect outcomes: ST elevation, oliguria, ventricular arrhythmias,... those variables should be included in the analysis. alternatively, limitations should

	be expanded as those factors were not present in the study and this might lead to reduced clinical effectiveness of this nomogram.

REVIEWER	Edward Koifman Soroka Medical Center, Ben-Gurion University of the Negev, Beer Sheva, Israel
REVIEW RETURNED	01-Aug-2020

GENERAL COMMENTS	The authors present a nomogram development for prediction of acute MI patients 30-day prognosis based on 5 parameters in a derivation population of 1422 patients and validation in 609 patients. The overall ROC was 0.765. Evaluating prognosis of acute MI patients is crucial for treatment decisions along with adjustment of ongoing therapy and therefore an accurate model is of significant value. My comment are as following: 1. Since lab tests change during the admission, which values were taken for the nomogram? 2. What is the performance of this nomogram compared to other ones in this cohort? what is the ROC for GRACE or TIMI etc. and the net reclassification index?

REVIEWER	Conrad Kabali Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Canada
REVIEW RETURNED	11-Aug-2020

GENERAL COMMENTS	The authors developed and validated a tool for predicting and classifying MI mortality in intensive care units. Overall, the method is sound, and the paper is well written. I have some comments. Page 2, line 9: The authors need to state that classification/discrimination is also part of the objective of the objective of the study Page 2, line 43: In my opinion, “excellent” is a strong word given that more than 20% of patients are likely to be misclassified. Okay to say, “reasonably accurate” (or something similar) Page 2, line 59: As in my previous comment, please replace “accurate” with something like “…a reasonable degree of accuracy”. Page 4, line 40: Please replace “multivariate” with “multivariable”. Also explain (briefly) how you checked for the proportional hazard assumptions Page 4, line 43: Performance of the model includes the ability of the model to discriminate. I think the term “discrimination capacity” is redundant and should be deleted Page 6, line 51: The authors need to explain how they determined that the missingness was not at random. Also, unclear if after imputation the covariates with less than 20% missing were treated as fixed or random. Put it another way, how uncertainty due to imputation was handled in the Cox model? Did they treat the imputed values as real in the prediction model? Page 7, line 30: Can the authors elaborate what they mean by “clinical usefulness” and how it was treated in the model Page 11, line 14: It seems like “hazard” was used to approximate risk because it is the term “risk” that is used by the authors to describe what the tool does. Can this be clarified in the Methods section? Also, a justification as to why “the hazard” is a good approximation of “30-day risk of dying” should be provided

REVIEWER	Guido Knapp TU Dortmund University Germany
REVIEW RETURNED	18-Aug-2020

GENERAL COMMENTS	Comments on statistics: The strategy for the statistical analyisis is clearly described. The methids used are appropriate. I have only one question: - Did the authors check the propotional hazard assumtipon? Please give at least a comment!

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Carmelo Dominici
Institution and Country: Università Campus Bio-Medico di Roma, Rome, Italy
Competing interests: None declared.

Please leave your comments for the authors below
Q1 authors should be congratulated for using a time-to-event analysis. however, considering that "30-days mortality" (a binary variable) was the primary endpoint, the rationale of choosing a Cox regression analysis should be described. proportional hazard assumption should be checked for each variable included in the (univariable) model when using Cox analysis. alternatively, logistic regression can be more easily performed to detect significant predictive variables.
Response: Thank you for your highly positive comments. In our study, the objective was to develop a fast-to-use prognostic model for 30-day mortality in critically ill myocardial infarction patients. And Cox proportional hazards model was the most frequently used regression model for survival analysis and thus was enrolled in this study. The rationale of choosing a Cox regression analysis was added in method (page 7).
The proportional hazards assumption was checked based on the scaled Schoenfeld residuals using survival package in R tool. The global model and all variable except heart rate were fitted for proportional hazards assumption. We have renewed method (page 7) and Supplementary Table 1 to show the result more clearly.

Q2 please discuss the rationale of the cut-off for inclusion in multivariable analysis. generally a P value of 0.1-0.2 in univariable analysis is used. from the manuscript, it appears that the cut-off was 0.05.
Response: Thank you for your valued suggestion. The final prognostic model was designed to be easily used in clinical work. A strict cut-off value of 0.05 was chosen to avoid too many variables entering into the final model and influencing the practicality of model. Details were added in method according to your valued suggestion (page 7).

Q3 survival curves should include 95% confidence intervals, number at risk for each group and censored cases. please include those 2 figures as separate images.
Response: Thank you for your valued suggestion. A new figure (Figure 3) was added. No censored cases were observed in this cohort. Confidence intervals and number at risk for each group were also added according to your suggestion.

Figure 3 Survival curves for two groups classified by prognostic total score calculated from the nomogram in the primary(A) and validation(B) cohort. For each survival curve, 95% confidence

intervals and number at risk for each group were also presented.

Q4 in results, please avoid text/table duplication. include most numeric results in tables to improve fluency of the manuscript.
Response: Thank you. Result part was revised according to your valued suggestion (page 8).

Q5 please provide 2 figures with examples of a low risk patient and a high risk patient (e.g. with numbers provided in page 9, lines 38-46).
Response: Thank you. Two new figures (Supplementary Figure 2 and Supplementary Figure 3) were added and the corresponding results (page 9) were also revised according to your valued suggestion.

Q6 differences in threshold probabilities between derivation cohort and validation cohort should be discussed.
Response: Thank you. A difference in threshold probability between primary and validation cohort was observed in our study. This difference may be due to the potential heterogeneity between these two cohorts, such as the level of variables or mortality rate, although which had not shown significant differences in statistical analyses. Overall, both two decision curves indicated a net benefit with respect to the use of nomogram model. Details were added in discussion part (page 12) according to your kind suggestion.

Q7 a table comparing this nomogram with existing models (e.g. number of patients included in cohorts, AUC, predictive probabilities, ...) should be provided.
Response: Thank you. A new table (Supplementary Table 2) and the corresponding description (page 11) were added according to your valued suggestion. Comparing with other existing models of which the AUC ranged from 0.66 to 0.90, the nomogram model showed an acceptable AUC of 0.80.

Q8 some factors of clinical presentation are known to affect outcomes: ST elevation, oliguria, ventricular arrhythmias, ... those variables should be included in the analysis. alternatively, limitations should be expanded as those factors were not present in the study and this might lead to reduced clinical effectiveness of this nomogram.
Response: Thank you. We felt sorry that these variables were not accessible in this study. Limitation part was expanded according to you valued suggestion.

Reviewer: 2

Reviewer Name: Edward Koifman
Institution and Country: Soroka Medical Center, Ben-Gurion University of the Negev, Beer Sheva, Israel
Competing interests: None declared

Please leave your comments for the authors below
The authors present a nomogram development for prediction of acute MI patients 30-day prognosis based on 5 parameters in a derivation population of 1422 patients and validation in 609 patients. The overall ROC was 0.765. Evaluating prognosis of acute MI patients is crucial for treatment decisions along with adjustment of ongoing therapy and therefore an accurate model is of significant value.
My comment are as following:
Q1 Since lab tests change during the admission, which values were taken for the nomogram?
Response: Thank you for your highly positive comment. Because we designed the nomogram model to be used as early as possible for patient, the first value in the initial 24 hours following admission was taken. We revised the method description (page 6) to claim it more clearly.

Q2 What is the performance of this nomogram compared to other ones in this cohort? what is the

ROC for GRACE or TIMI etc. and the net reclassification index?
Response: Thank you for your kind comment. We felt sorry that GRACE or TIMI score were not accessible in this study and thus the comparison could not made. A new table (Supplementary Table 2) were added to present AUC of different model in other studies. Comparing with other existing models of which the AUC ranged from 0.66 to 0.90, the nomogram model showed an acceptable AUC of 0.80. In addition, limitation part (page 13) was revised to point out this limitation more clearly.

Reviewer: 3

Reviewer Name: Conrad Kabali
Institution and Country: Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Canada
Competing interests: None declared

Please leave your comments for the authors below
The authors developed and validated a tool for predicting and classifying MI mortality in intensive care units. Overall, the method is sound, and the paper is well written. I have some comments.

Q1 Page 2, line 9: The authors need to state that classification/discrimination is also part of the objective of the objective of the study
Response: Thank you for your highly positive comment. The objective was revise according to your kind suggestion.

Q2 Page 2, line 43: In my opinion, “excellent” is a strong word given that more than 20% of patients are likely to be misclassified. Okay to say, “reasonably accurate” (or something similar)
Response: Thank you. The word “excellent” was replaced with “reasonably accurate” according to your kind suggestion.

Q3 Page 2, line 59: As in my previous comment, please replace “accurate” with something like “…a reasonable degree of accuracy”.
Response: Thank you for your kind suggestion. The word “accurate” was replaced with “achieve a reasonable degree of accuracy to predict…”.

Q4 Page 4, line 40: Please replace “multivariate” with “multivariable”. Also explain (briefly) how you checked for the proportional hazard assumptions
Response: Thank you. All “multivariate” were replaced with “multivariable”. The proportional hazards assumption was checked based on the scaled Schoenfeld residuals using survival package in R tool. We have renewed method (page 7) and Supplementary Table 1 to show the result more clearly.

Q5 Page 4, line 43: Performance of the model includes the ability of the model to discriminate. I think the term “discrimination capacity” is redundant and should be deleted
Response: Thank you. We do agree with your comment and deleted this term.

Q6 Page 6, line 51: The authors need to explain how they determined that the missingness was not at random. Also, unclear if after imputation the covariates with less than 20% missing were treated as fixed or random. Put it another way, how uncertainty due to imputation was handled in the Cox model? Did they treat the imputed values as real in the prediction model?
Response: Thank you for your comment. Data in this study was derived from MIMIC database, which was collected in critical care unit. We hypothesized that the missingness of certain variable was due to the judgement of clinicians. That’s to say, clinicians determined whether one examination was carried out or not mainly based on their judgement on the patient status but not at random. Thus, multiple imputation was used by reference to some previous studies [1-2].

To make it as simplified as possible, one dataset in which the value of coefficient was the nearest to the pooled value was shown. In this dataset, the covariates after imputation were treated as fixed and results were calculated based on this dataset.

[1] Zhang, Z., Mo, L., Ho, K.M., and Hong, Y. (2019). Association Between the Use of Sodium Bicarbonate and Mortality in Acute Kidney Injury Using Marginal Structural Cox Model. Crit Care Med 47, 1402-1408.
[2]Shahn, Z., Shapiro, N.I., Tyler, P.D., Talmor, D., and Lehman, L.-W.H. (2020). Fluid-limiting treatment strategies among sepsis patients in the ICU: a retrospective causal analysis. Crit Care 24, 62.

Q7 Page 7, line 30: Can the authors elaborate what they mean by “clinical usefulness” and how it was treated in the model
Response: Thank you for your comment. Clinical usefulness was defined as the ability to make better decisions with a model than without. In our study, decision curve analysis was used to evaluate the clinical usefulness of nomogram model. According to decision curve analysis, the use of nomogram model could achieve net benefit under a relatively wide range of event probability.

Q8 Page 11, line 14: It seems like “hazard” was used to approximate risk because it is the term “risk” that is used by the authors to describe what the tool does. Can this be clarified in the Methods section? Also, a justification as to why “the hazard” is a good approximation of “30-day risk of dying” should be provided
Response: Thank you for your valued suggestion. We do agree that the term “hazard” calculated by Cox regression was used to approximate risk. The description was clarified in Methods section in revised manuscript (page 7). Death event may occur across a range of time period, and hazard was considered as a summary estimate of the death event in survival data. That’s to say, “hazard” represented the death event rate for an individual under the time period of “30-day” and thus was used as an approximation of “30-day risk of dying” in our study.

Reviewer: 4

Reviewer Name: Guido Knapp
Institution and Country:
TU Dortmund University
Germany
Competing interests: None declared!

Please leave your comments for the authors below
Comments on statistics:

The strategy for the statistical analyisis is clearly described. The methids used are appropriate. I have only one question:
- Did the authors check the propotional hazard assumtipon? Please give at least a comment!
Response: Thank you for highly positive comment. The proportional hazards assumption was checked based on the scaled Schoenfeld residuals using survival package in R tool. We have renewed method (page 7) and Supplementary Table 1 to show the result more clearly.

VERSION 2 – REVIEW

REVIEWER	Carmelo Dominici Università Campus Bio-Medico di Roma, Rome, Italy
REVIEW RETURNED	08-Oct-2020

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.

REVIEWER	Edward Koifman Soroka Medical Center, Israel
REVIEW RETURNED	31-Oct-2020

GENERAL COMMENTS	There are parameters that their univariate analysis is not shown such as pH and lactate which could also be of use for the model Please provide univariate report for all sucha parameters

REVIEWER	Conrad Kabali Division of Epidemiology Dalla Lana School of Public Health University of Toronto Canada
REVIEW RETURNED	15-Oct-2020

GENERAL COMMENTS	All comments have been addressed. No further comments.

REVIEWER	Guido Knapp Department of Statistics TU Dortmund University Germany
REVIEW RETURNED	12-Oct-2020

GENERAL COMMENTS	Comments on Statisitcs: No further comments. Thank you!