PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and validation of dynamic models to predict post-discharge mortality risk in patients with acute myocardial infarction: results from China Acute Myocardial Infarction registry
AUTHORS	Lv, Junxing; Wang, Chuang-Shi; Gao, Xiaojin; Yang, Jingang; Zhang, Xuan; Ye, Yunqing; Dong, Qiuting; Fu, Rui; sun, hui; Yan, Xinxin; Zhao, Yanyan; Yang, Wang; Xu, Haiyan; Yang, Yuejin

VERSION 1 – REVIEW

REVIEWER	Rinaldi, Riccardo Catholic University of the Sacred Heart
REVIEW RETURNED	12-Dec-2022

GENERAL COMMENTS	The article is nicely written and quite informative. The sample size is conspicuous, the statistical analysis is well conducted, and the results are interesting. I congratulate the authors for their study. Here are some comments for the authors that could improve the value of the manuscript: 1) The authors reported that all types of AMI were eligible for the CAMI registry, except type 4a and type 5. Given the different pathogenesis and prognosis associated with the type of MI, the authors should also report, if available, the percentage of the different group of MI (i.e., type I, type II and type III) in the overall population to show if one type is more represented than another, as this could have biased the results. 2) Are any angiographic features available? Indeed, the number of diseased vessels, the culprit vessel as well as the localization (proximal vs. distal) or other angiographic features (e.g., intraprocedural complications) can have a relevance in influencing the prognosis. If not, the authors should acknowledge this as a limitation of the study. 3) I would like the authors to further implement the Discussion section by quickly discussing the clinical and therapeutic implications of their nomograms. Similarly, they could implement the discussion further highlighting the difference from the previous nomograms as well as the associated clinical and therapeutic implications. 4) An important limitation of this study is obviously its retrospective nature, and this should be acknowledged in the limitations section. 5) Throughout editing by a native English speaker could improve the readability of the manuscript. 6) Please double check all abbreviations and expand them at their first use.

REVIEWER	Luney, Matthew University of Nottingham, Anaesthesia & Critical Care
REVIEW RETURNED	23-Dec-2022

GENERAL COMMENTS	I thank the authors for this interesting and well written manuscript. I have the attached comments including queries to address for revision (mostly methodologic in nature) for consideration before supporting publication. Is the article important? Yes. The burden of cardiovascular disease is common and prognosticating after acute myocardial infarction (AMI) is important. Whilst other prognostic models for survival after AMI exist, this is the most up-to-date model for this particular context of AMI surivors in China in 2013–14. This manuscript presents two internally validated prognostic models of survival (Model 1 surivival to 30 days and Model 2 survival to two years) derived from a prospectively constructed registry of survivors to hospital discharge after AMI in 108 hospitals in mainland China between 2013 and 2014 who were followed up for 2 years. Will it help our readers make better decisions and, if so how? Yes. For healthcare providers looking after survivors of AMI the authors' models provide useful prognostic information which they can share with their patients. As the authors rightly observe their model requires external validation before it is can be used in other settings, such as settings where PCI use may be higher (it was 57% as per Table 1) or where the prevalence of current smokers is lower than 45% (Table 1). This manuscript's main output – the prognostic models will be of most use in contexts with similar healthcare resource use and comparable population health. Given the large geographic and population coverage of their registry this will be applicable to many patients/future survivors of AMI. Will the article add enough to existing knowledge? Yes. The authors assert their prognostic models significantly outperform existing models, which is correct with respect to statistically significant improvements in IDI, BIC, p-value and LR ratio test. The main strength of this article is that the authors that they extend their model to 2 years as most work has only reported prognostication to 30 days post AMI (except for GRACE2.0 which now reaches 3 years). This model is novel for the authors context, there are limited existing publications regarding risks after AMI in their setting, and those that do exist are risk factors identified in this CAMI registry or similar national registries. This article offers fresh insights into the post discharge risks in a healthcare setting where the provision of PCI for AMI is continuing to increase and as such existing data on survivorship

	continues to improve. However some risk prediction models after AMI do already exist such as: doi.org/10.1136/bmj.38985.646481.55 doi.org/10.1016/j.amjcard.2016.07.029 doi.org/10.1136/heartjnl-2016-309359 doi.org/10.1016/j.jjcc.2021.06.002 And risk factors are already published such as doi.org/10.1001/jamanetworkopen.2018.1079 Many of the article's findings on predictors are confirmation of already recognised risk factors, including the authors' own work previously published in this journal (doi.org/10.1136/bmjopen2019-030772) In developing the models the authors found an association between taking statins (doi.org/10.1016/S0140-6736(05)67394-1) or beta-blockers (doi.org/10.1161/CIRCULATIONAHA.116.026336) and increased survival after hospital discharge following AMI. However these are not novel findings, indeed it has been internationally accepted guidance for decades to treat patients with these medications after discharge: doi.org/10.1161/01.CIR.94.9.2341 doi.org/10.1161/CIR.0b013e3182742cf6 doi.org/10.1093/eurheartj/ehx393 https://www.nice.org.uk/guidance/ng185 Does the article read well and make sense? Does it have a clear message? The article is well written, coherent and conveys clearly the objective and the outputs. PPI It is disappointing that despite enrolling over 20,000 patients in the authors registry, there were no patients or lay public involved in the research to establish what risks are important to people who survive to hospital discharge after AMI. Whilst mortality incontrovertibly an undesirable outcome, it is not a particularly patient centred outcome – such as readmission free time, repeat infarction, quality of life (e.g. EQ5D). Is the research question clearly defined and appropriately answered? Yes Overall design of study - appropriate and adequate to answer the research question? Yes Methods comments A strength was the authors consideration of collinearity, although they do not actually report what their findings were in this regard. There was also no discussion about interaction terms. For instance, the authors use several measures related to heart failure: Killip class, LVEF, HF during hospitalization, HF working 30 days after discharge, and they also use discharge on beta-blocker therapy. These predictors almost certainly interact with each other. The time dependent nature of risk, which is essentially what is implied by the claim of dynamic risk prediction is essentially dichotomised into risk at 30 days and risk at 2 years.

	Authors appear to have chosen their predictors from what they had available in the registry and then pared this down using LASSO. Their application of LASSO appears sound. The authors of this manuscript have indeed themselves already published on this resgirary cohort demonstrating AF is assocaited with inhospital mortality after AMI but not then used this in their model. Their model also includes predictors that are very similar such as Killip class, In-hospital HF and 30- day HF worsening, LVEF – from a clinical perspective these are sufficiently closely related that any neglible statistical gains are outweighed be the additional burden collecting each predictor to use in a prognostic model. A more parsimonious approach would be preferable in clinical practice. A technical point these models are not dynamic, they are two individual models each informing probability of survival to a specific time point. The authors essentially present two separate models that prognosticate 30 day survival after hospital discharge and 2 year survival after hospital discharge. They do not for instance give any 'dynamic' information on likelihood of survival between these two time points other than that it will of course lie between the 30 day and 2 year probabilities. More flexible techniques such as spline adjustment or landmarking could have achieved a more dynamic model, or use of generalised additive model as opposed to the cox regression. Interpretation and conclusions The interpretation and conclusions are sound References These are appropriate Abstract The abstract is well written and accurately reflects the contents of the full manuscript. Revisions/Queries to answer: - It is not clear to me why the authors applied the older GRACE from 2004 not the updated GRACE 2.0 predictive model from 2014 (also published in this journal, doi.org/10.1136/bmjopen-2013-004425)? GRACE 2.0 was developed in additional countries and unlike the original GRACE this included China – no countries in Asia were represented in the original GRACE model. - Why did the authors choose to compare the models at 30 days and 2 years when GRACE is a prognostic model for in hospital mortality (these patients were excluded in this article) and at 6 months? - Perhaps if the co-efficients of GRACE2.0 were used this would be a more fair comparison or justification as to why this was not done. - The authors make specific mention that events during the admission or shortly after

	discharge are of important prognostic value e.g. recurrent myocardial ischaemic, HF worsening within 30 days: Why do they not discuss important lifestyle interventions (e.g. smoking cessation doi.org/10.1161/CIRCULATIONAHA.109.891523), or rehabilitation (e.g. cardiac rehabilitation programmes doi.org/10.1016/j.jacc.2015.10.044) which are incredibly well evidenced yet not factored in to their model? Equalty, diversity and inclusivity Why have the authors not discussed ethinicity/socioeconomic data despite these being known risk factors for survival after myocardial infarction? doi.org/10.1371/journal.pone.0065130 doi.org/10.1186/s12963-021-00280-1 doi.org 10.1016/j.gaceta.2011.06.013 doi.org/10.1161/CIRCULATIONAHA.115.017009 Methodology Prospective? The data collection was done prospectively but this particular article is not a prospective study – the research question and the candidate predictors were identified after the data collection which occurred between January 2013 to September 2014. The model co-variates were not pre-specified prior to data collection. This is fine but this paper is not a prospective study, albeit the data were collected prospectively this is secondary analysis of a registry. Predictors Why is an assessment of linearity of predictors not mentioned. Yes there is mention of an assessment of VIF with respect to multicollinearity but this is a between predictor phenomenon. Authors should have assessed whether their predictors (such as age) should have been handled as linear predictors, or if they have already done so this should be stated. With respect to their cox regression – did the authors plot the Schoenfield residuals to test the proportionality assumption? Why was hospital level not used as a predictor? In previously published work for the registry data collection (doi.org/10.1001/jamanetworkopen.2020.21677) the authors demonstrate the CAMI registry data are collected from three different hospital types (provence, prefecture and county). They go on to describe in detail how the outcomes when treated for AMI in these three difference hospital types is associated with wide variation in mortality after AMI: 3.1%, 5.3% and 10.2% respectively. These are starkly different in-hospital mortality rates and it is reasonable to expect with such wide variation to persist after discharge. However the authors do not use hospital level in their regression model as a predictor, importantly though there is also no discussion about why they have not done so. I do not mention it to mandate that this is a predictor, although from their

	previous work one can see they have this data available, but that I would strongly recommend analysis that either demonstrates this predictor is justifiably excluded despite clear evidence from their own work that it is a strong risk factor for mortality after AMI. Missing data? Their subgroup analyses in Supplemental Tables 4 and 5 only have approx. 5600 and 4400 patients each. These subgroups were for age, sex, diabetes, AMI classification and PCI. These are major components of their main prognostic models, I am concerned that there was a large amount of missingness in their dataset if for instance they were only able to perform subgroup analysis by age on 5701 of their 23887 patients. Indeed this appears to suggest there is considerable missingness in their main analysis which is superficially addressed in the main text where they refer to performing the MICE. The chosen strategy of multiple imputation itself is not unreasonable and it's execution appears sound. However, with respect to the TRIPOD statement item 13b "including the number of participants with missing data for predictors and outcome" they do not report this adequately for each predictor. They only report missingness for 30 day medication use. I note table 1 is reported without imputated data which is appropriate. However the proportion of missingness should also be reported – for instance if the missingness for age is up to 76% of age data missing this is of signficant concern. I would strongly recommend authors to openly publish the level of missing data for each predictor used at least in the final models. Why have the authors not commented on the very low rate of COPD? The rate of COPD is surprisingly low in this cohort (<1.9%) despite current smoker rate of c45% but no comment on this from the authors. Cf UK population of AMI the incidence of COPD is 12% (doi.org/10.1136/heartjnl-2016-309860). The incidence of prior heart failure is also surprisingly low (only c2%) Time dependent co-variates? Is it truly a dynamic model – the authors prognosticate for 30 day and 2 year survival after AMI. They do not dynamically prognosticate in so far as one can only interpret probability of survival to two specific timepoints. In a dynamic model one would expect to be able to ascertain prognosis at any timepoint after the event for instance what about at 1 year after AMI? The model could be strengthen by including the variable "time since AMI" as a predictor. It is impossible to get away from survivor bias but in this prognostic model that is relevant – those who survive to 1 year after AMI are not the same as those who have only survived 30 days after AMI but they are both represented the same in the 2 year prognostic model. Other minor points

	Table 1 - A minor point but inferential statistics such as a p-value in Table 1 is of limited to no use given the authors have already stated that they randomly allocated patients to their derivation or validation cohorts. There should be no reason to expect imbalanced groups unless failure of randomization. This is also true for Supplemental Table 1. - Age when presented in years should be a whole number unless the authors recorded the age in greater detail than whole years. - Heart rate should be reported in whole beats per minute. Spelling and Grammar – minor amendments On line 22 page 12 it should read “than GRACE score” not “CRACE” On line 21 page 18 it should read “informed consent was given” not “consents were” The titles for Supplemental Figures 16 and 17 should begin: “Comparison of Clinical Utility…”, the plural of utility in this context is still utility. Same for the use of the word utility in the caption text for Supplemental Fig 16 & 17. The same is true of Line 9 page 12 it should read “better clinical utility” Regarding Tripod checklist Item 1: Title does not state the risk being predicted: Mortality Item 1: Title does not state that the population is patients who had already survived to hospital discharge Item 8: Authors state the sample size but have not performed a sample size calculation. i.e. no justification obtaining sufficient events per predictor used in their model. A crude calculation from their 16 predictors (accounting for the multiple levels within their variables such as Killip classes) and 190 events in their 30 day model there are 11.8 events per predictor). It would be good practice for authors to show their sample size calculation/justify their sample size rather than just use all the data available. Item 13: significant concerns about lack of missing data reporting. How it was handled (Item 9) is reported (multiple imputation) however the reason which in the case of missing prescribing data cases were excluded instead of imputed is not justified. Item 18: Missingness is not discussed in the limitations particularly with respect to the variables used as predictors

VERSION 1 – AUTHOR RESPONSE

To reviewer 1:

We are grateful for the reviewer’s careful review, important comments, and useful suggestions. Our responses are presented under each comment. Revisions were carefully made in the revised version of both the manuscript and online supplemental materials.

1) The authors reported that all types of AMI were eligible for the CAMI registry, except type 4a and type 5. Given the different pathogenesis and prognosis associated with the type of MI, the authors should also report, if available, the percentage of the different group of MI (i.e., type I, type II and type III) in the overall population to show if one type is more represented than another, as this could have biased the results.

Response: Thank you for this insightful comment. We totally agree that if the distribution of types of AMI was unbalanced, it could have biased the results of the present study. Details of types of AMI were not collected in the CAMI registry, so the distribution was not presented. However, the CAMI registry enrolled AMI patients consecutively from 108 hospitals throughout Mainland China, which meant that it was representative of AMI population in routine clinical practice. The impact of distribution of AMI types might be relatively limited. We further added this content in the limitation section (**Page 19, Line 6-11**), as below.

"Third, the distribution of AMI types (types 1, 2, 3, 4b, and 4c) was not collected in the CAMI registry. Results of the present study could have biased if a certain type was more represented than another. However, the CAMI registry enrolled patients consecutively from 108 hospitals, which meant that it was representative of AMI population in routine clinical practice. It is plausible that the impact of distribution of AMI types is relatively limited."

2) Are any angiographic features available? Indeed, the number of diseased vessels, the culprit vessel as well as the localization (proximal vs. distal) or other angiographic features (e.g., intraprocedural complications) can have a relevance in influencing the prognosis. If not, the authors should acknowledge this as a limitation of the study.

Response: We are grateful for this comment. We agree that including the angiographic features may further improve risk prediction in patients following AMI. Unfortunately, angiographic features such as the localization were not available in a large proportion of the present cohort. We acknowledged this as a limitation of the study (**Page 18, Line 19-22**), as below.

"Although the present risk prediction tool has achieved satisfying discrimination and calibration, it may be further improved by including other prognostic factors of AMI, such as details of angiographic characteristics, which were not available in a large proportion of the cohort."

3) I would like the authors to further implement the Discussion section by quickly discussing the clinical and therapeutic implications of their nomograms. Similarly, they could implement the discussion further highlighting the difference from the previous nomograms as well as the associated clinical and therapeutic implications.

Response: Thank you very much for this helpful suggestion. We further added discussion to highlight the differences of the present prognostic nomograms with previous nomograms as well as the relevant clinical and therapeutic implications (**Page 18, Line 2-14**), as below.

"Prognostic nomogram is a graphical presentation format for complex predictive regression model.[33] A series of prognostic nomograms have been established for risk prediction in patients with cancer or cardiovascular diseases.[34-39] For patients with myocardial infarction, previous prognostic nomograms mainly focused on evaluating short-term risk of mortality or other adverse events.[37,38] There also existed nomogram developed to predict risk of adverse events beyond 1 year.[36] However, without consideration of changing nature of event risk or medications, the nomogram might not play roles in post-discharge management of patients. Our prognostic nomograms, which took into account follow-up adverse event as well as medications, could assist in risk reassessment at 30 days after discharge. In detail, using the nomogram for prediction of 30-day mortality, physicians can identify high-risk patients at discharge. At 30-day follow up, the second nomogram can be used to reassess mortality risk of 30-day survivors, and may guide decision-making of long-term follow-up intensity and strategies of medical care."

4) An important limitation of this study is obviously its retrospective nature, and this should be acknowledged in the limitations section.

Response: We appreciate this constructive comment. We added the content in the limitation section (**Page 18, Line 17-19**).

"First, as a retrospective analysis of a prospective cohort, this study only used data which had been collected in the CAMI registry.."

5) Throughout editing by a native English speaker could improve the readability of the manuscript.

Response: Thank you for this helpful suggestion. We have carefully checked and revised the manuscript to improve the readability.

6) Please double check all abbreviations and expand them at their first use.

Response: Thank you for this helpful suggestion. We have double checked and revised the abbreviations in the manuscript.

To reviewer 2:

We express our great gratitude for the reviewer's careful review, constructive comments, and helpful suggestions. Our responses are presented under each comment, and revisions were carefully made in the revised version of both the manuscript and online supplemental materials.

1) It is not clear to me why the authors applied the older GRACE from 2004 not the updated GRACE 2.0 predictive model from 2014 (also published in this journal, doi.org/10.1136/bmjopen-2013-004425)? GRACE 2.0 was developed in additional countries and unlike the original GRACE this included China– no countries in Asia were represented in the original GRACE model.

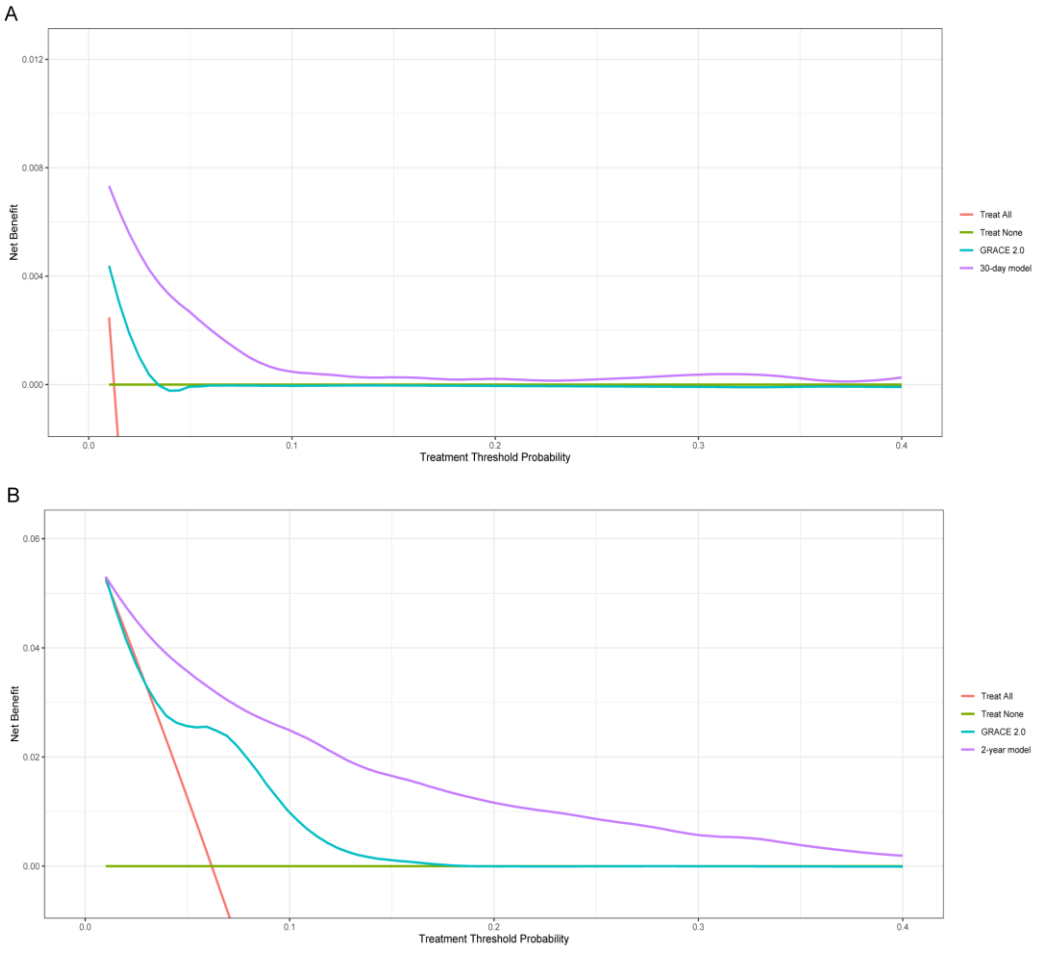
Response: Thank you very much for this helpful comment. We further compared the predictive performance between our risk prediction models and GRACE 2.0 score (doi: 10.1136/bmjopen-2013-004425). Both 30-day and 2-year models showed significantly better predictive performance than the GRACE 2.0 score (30-day model vs GRACE 2.0 score: C index, 0.855 [0.830-0.879] vs 0.752 [0.720-0.784]; NRI [95%CI], 0.569 [0.500-0.624], P<0.0001; IDI [95%CI], 0.061 [0.044-0.101], P<0.0001; BIC, 3247.357 vs 3492.004; 2-year model vs GRACE 2.0 score: C index, 0.825 [0.811-0.839] vs 0.769 [0.752-0.786]; NRI [95%CI], 0.486 [0.456-0.529], P<0.0001; IDI [95%CI], 0.115 [0.098-0.143], P<0.0001; BIC, 12257.375 vs 12934.783; **Page 13, Line 6-8, Line 12-14; online supplemental figure 23**).

"For predicting 30-day mortality, the 30-day risk prediction model showed significantly better predictive performance than both GRACE 1.0 and 2.0 scores (30-day risk model vs GRACE 1.0 score: C index, 0.855 [0.830-0.879] vs 0.771 [0.740-0.802]; NRI [95%CI], 0.412 [0.307-0.485],

P<0.0001; IDI [95%CI], 0.048 [0.032-0.090], P<0.0001; BIC, 3267.271 vs 3402.578; 30-day risk model vs GRACE 2.0 score: C index, 0.855 [0.830-0.879] vs 0.752 [0.720-0.784]; NRI [95%CI], 0.569 [0.500-0.624], P<0.0001; IDI [95%CI], 0.061 [0.044-0.101], P<0.0001; BIC, 3247.357 vs 3492.004).”

“Similarly, when predicting 2-year mortality, the 2-year risk prediction model also performed better than the GRACE risk scores (2-year risk model vs GRACE 1.0 score: C index, 0.825 [0.811-0.839] vs 0.798 [0.783-0.813]; NRI [95%CI], 0.191 [0.147-0.257], P<0.0001; IDI [95%CI], 0.041 [0.031-0.057], P<0.0001; BIC, 12540.559 vs 12697.527; 2-year risk model vs GRACE 2.0 score: C index, 0.825 [0.811-0.839] vs 0.769 [0.752-0.786]; NRI [95%CI], 0.486 [0.456-0.529], P<0.0001; IDI [95%CI], 0.115 [0.098-0.143], P<0.0001; BIC, 12257.375 vs 12934.783). The decision curve analysis further demonstrated better clinical utility of both 30-day and 2-year risk models than GRACE scores (online supplemental figure 22 and 23).”

*“**Supplemental Figure 23. Comparisons of clinical utility between models and GRACE 2.0 score***



”

2) Why did the authors choose to compare the models at 30 days and 2 years when GRACE is a prognostic model for in hospital mortality (these patients were excluded in this article) and at 6 months?

Response: Thank you for raising this question. We chose to compare the models with the GRACE risk score because it was a validated prognostic tool which had been already applied into routine clinical practice. Although the GRACE score was developed for estimating the risk of in-hospital or 6-month post-discharge mortality, previous studies have showed that it had good predictive value for 30-day and 2-year mortality (doi: 10.2147/VHRM.S117204, doi: 10.1016/j.ijcard.2012.09.076). The comparisons between our models and the GRACE 2.0 score (doi: 10.1136/bmjopen-2013-004425) were further added in the manuscript (**Page 13, Line 6-8, Line 12-14; online supplemental figure 23**).

3) Perhaps if the co-efficients of GRACE 2.0 were used this would be a more fair comparison or justification as to why this was not done.

Response: We appreciate this comment. GRACE 2.0 score is a validated prognostic tool in patients with myocardial infarction, and we further added the comparisons between the present models and the GRACE 2.0 score (**Page 13, Line 6-8, Line 12-14; online supplemental figure 23**).

4) The authors make specific mention that events during the admission or shortly after discharge are of important prognostic value e.g. recurrent myocardial ischaemic, HF worsening within 30 days: Why do they not discuss important lifestyle interventions (e.g. smoking cessation doi.org/10.1161/CIRCULATIONAHA.109.891523), or rehabilitation (e.g. cardiac rehabilitation programmes doi.org/10.1016/j.jacc.2015.10.044) which are incredibly well evidenced yet not factored in to their model?

Response: Thank you for raising this question. We agree that the inclusion of important lifestyle interventions and rehabilitation may further improve risk prediction of patients following AMI. Unfortunately, these data were not collected in the CAMI registry. We realize that this is an important limitation of this study, and added the relevant content in the limitation section (**Page 18, Line 23-25; Page 19, Line 1**), as below.

"However, lifestyle interventions and cardiac rehabilitation programmes, which were associated with lower risk of adverse events in patients with coronary artery disease,[40,41] as well as laboratory and echocardiographic indexes were not collected during follow up."

5) Why have the authors not discussed ethnicity/socioeconomic data despite these being known risk factors for survival after myocardial infarction?
doi.org/10.1371/journal.pone.0065130

doi.org/10.1186/s12963-021-00280-1

doi.org/10.1016/j.gaceta.2011.06.013 doi.org/10.1161/CIRCULATIONAHA.115.017009

Response: We appreciate this helpful comment. We further added the relevant contents in the discussion and limitation sections (**Page 17, Line 13-17; Page 19, Line 12-14**), as below.

"Socioeconomic factors, which were known as risk factors for survival following myocardial infarction,[26-29] were not included in the present models because we sought to develop models based on predictors directly reflecting patients' clinical conditions. Notably, these factors were also not included in existing risk prediction tools.[6,7,9,16]"

"Finally, our dynamic models were only internally validated in Chinese patients. Further validations in external cohorts including patients of other races are needed."

6) **Prospective?** The data collection was done prospectively but this particular article is not a prospective study – the research question and the candidate predictors were identified after the data collection which occurred between January 2013 to September 2014. The model covariates were not pre-specified prior to data collection. This is fine but this paper is not a prospective study, albeit the data were collected prospectively this is secondary analysis of a registry.

Response: Thank you for raising this important issue. We have accordingly revised the abstract and added the content in the limitation section (**Page 3, Line 7; Page 18, Line 17-21**), as below.

*"**Design:** A retrospective analysis of a prospective cohort. "*

"First, as a retrospective analysis of a prospective cohort, this study only used data which had been collected in the CAMI registry. Although the present risk prediction tool has achieved satisfying

discrimination and calibration, it may be further improved by including other prognostic factors of AMI..."

7) **Predictors.** Why is an assessment of linearity of predictors not mentioned. Yes there is mention of an assessment of VIF with respect to multicollinearity but this is a between predictor phenomenon. Authors should have assessed whether their predictors (such as age) should have been handled as linear predictors, or if they have already done so this should be stated.

Response: Thank you for raising this question. Before regression analysis, we had checked the linearity assumption by the Martingale residual plots. We further added the relevant contents in the method section as well as supplemental materials (**Page 10, Line 14 and 15; online supplemental figure 2 and 3**), as below.

"Before regression analysis, we used Martingale residual plots to check the linearity assumption for continuous variables (online supplemental figure 2 and 3)."

8) With respect to their cox regression – did the authors plot the Schoenfield residuals to test the proportionality assumption?

Response: Thank you for raising this question. Before regression analysis, we had examined the proportional hazards assumptions by inspection of Schoenfeld residual plots. The contents were further added in the method section and supplemental materials (**Page 10, Line 16-18; online supplemental figure 4 and 5**), as below.

"The proportional hazards assumptions were tested by inspection of Schoenfeld residual plots (online supplemental figure 4 and 5)."

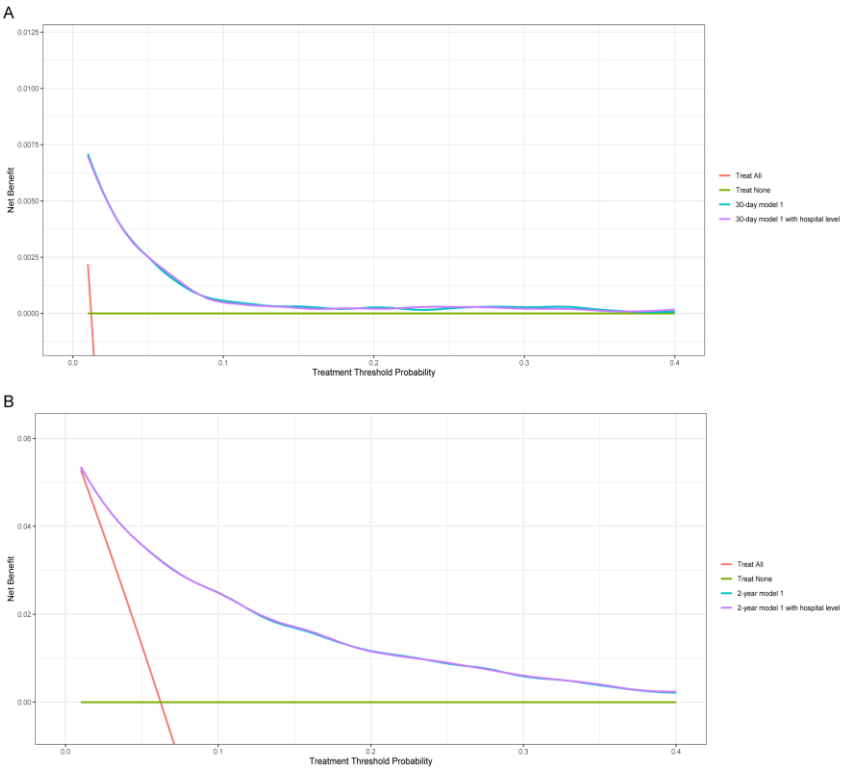
9) **Why was hospital level not used as a predictor?** In previously published work for the registry data collection (doi.org/10.1001/jamanetworkopen.2020.21677) the authors demonstrate the CAMI registry data are collected from three different hospital types (provence, prefecture and county). They go on to describe in detail how the outcomes when treated for AMI in these three difference hospital types is associated with wide variation in mortality after AMI: 3.1%, 5.3% and 10.2% respectively. These are starkly different in-hospital mortality rates and it is reasonable to expect with such wide variation to

persist after discharge. However the authors do not use hospital level in their regression model as a predictor, importantly though there is also no discussion about why they have not done so. I do not mention it to mandate that this is a predictor, although from their previous work one can see they have this data available, but that I would strongly recommend analysis that either demonstrates this predictor is justifiably excluded despite clear evidence from their own work that it is a strong risk factor for mortality after AMI.

Response: We express our great gratitude for this insightful comment which is important to improve this study. We have further added analyses and discussions about the hospital level. Considering that the improvement of care quality in relatively low-level hospitals in recent years was likely to weaken its prognostic value, patient-level characteristics were mainly taken into account for prediction. Additionally, we analyzed the potential prognostic value of hospital level beyond the current model predictors, and found that the hospital level provided no incremental prognostic value to 30-day or 2-year risk models (**Page 12, Line 21-23; online supplemental figure 21**). An additional subgroup analysis was also performed according to hospital level, and showed that both 30-day and 2-year models achieved satisfying predictive performance regardless of hospital levels (**online supplemental table 5 and 6**). Moreover, we discussed these results as well as the reason why we did not use hospital level in the regression model as a predictor (**Page 17, Line 8-13**), as below.

"Notably, the hospital level provided no incremental value to 30-day or 2-year risk models (the inclusion of hospital level to 30-day model, likelihood ratio test P=0.4174; to 2-year model, likelihood ratio test P=0.5621; online supplemental figure 21)."

"Supplemental Figure 21. Comparisons of clinical utility between models with or without hospital level



”

“Although a previous study from CAMI registry showed that there were significant variations in in-hospital mortality among three levels of hospitals in China,[14] hospital level was not used as a predictive index in the present risk prediction models, for the improvement of care quality in relatively low-level hospitals was likely to weaken its prognostic value. Besides, the hospital level was showed to provide no additional prognostic information beyond current predictors in the risk prediction models.”

10) **Missing data?**

Their subgroup analyses in Supplemental Tables 4 and 5 only have approx. 5600 and 4400 patients each. These subgroups were for age, sex, diabetes, AMI classification and PCI. These are major components of their main prognostic models, I am concerned that there was a large amount of missingness in their dataset if for instance they were only able to perform subgroup analysis by age on 5701 of their 23887 patients.

Indeed this appears to suggest there is considerable missingness in their main analysis which is superficially addressed in the main text where they refer to performing the MICE. The chosen strategy of multiple imputation itself is not unreasonable and it’s execution appears sound. However, with respect to the TRIPOD statement item 13b “including the number of participants with missing data for predictors and outcome” they do not report this adequately for each predictor. They only report missingness for 30 day medication use. I note table 1 is reported without imputated data which is

appropriate. However the proportion of missingness should also be reported – for instance if the missingness for age is up to 76% of age data missing this is of signficant concern. I would strongly recommend authors to openly publish the level of missing data for each predictor used at least in the final models.

Response: We appreciate this comment. Firstly, we would like to clarify that the subgroup analyses (**online supplemental table 5 and 6**) were performed in patients in the validation cohorts with complete data on model predictors instead of the total study population (n=23887). We have revised the relevant content in the method section to make it clearer (**Page 10, Line 11-13, Line 20 and 21**). Additionally, as suggested, number of missing values for predictors in the risk prediction models were further summarized in **online supplemental table 1**, as below. Only LVEF had >10% missing values in the total study population.

"Subgroup analyses were performed in patients with complete data on model predictors in the validation cohort according to age, sex, diabetes, AMI classification, in-hospital PCI, and hospital level (province level, prefecture level, and county level)."

"Number of missing values for selected predictors were shown in online supplemental table 1."

*"**Supplemental Table 1. Number of missing values for selected predictors in derivation and validation cohorts***

	Number of missing values (%)
Derivation cohort	
30-day prognostic model	
Age	*336 (2.1)*
Prior stroke	*820 (5.1)*
Heart rate	*391 (2.5)*
Killip class	*395 (2.5)*
LVEF	*3370 (21.2)*
In-hospital PCI	*435 (2.7)*
In-hospital recurrent myocardial ischemia	*518 (3.3)*
In-hospital recurrent myocardial infarction	*518 (3.3)*
In-hospital heart failure	*502 (3.2)*
Antiplatelet therapy at discharge	*849 (5.3)*
Statins at discharge	*849 (5.3)*

2-year prognostic model	
Age	*242 (2.0)*
Prior renal dysfunction	*510 (4.2)*
History of heart failure	*490 (4.0)*
AMI classification	*0 (0.0)*
Heart rate	*107 (0.9)*
Killip class	*114 (0.9)*
Hemoglobin	*327 (2.7)*
LVEF	*2307 (19.0)*
In-hospital PCI	*130 (1.1)*
In-hospital heart failure	*155 (1.3)*
Heart failure worsening within 30 days	*8 (0.1)*
Antiplatelet therapy within 30 days	*0 (0.0)*
β blockers within 30 days	*0 (0.0)*
Statins within 30 days	*0 (0.0)*
Validation cohort	
30-day prognostic model	
Age	*144 (1.8)*
Prior stroke	*404 (5.1)*
Heart rate	*197 (2.5)*
Killip class	*176 (2.2)*
LVEF	*1678 (21.1)*
In-hospital PCI	*216 (2.7)*
In-hospital recurrent myocardial ischemia	*267 (3.4)*
In-hospital recurrent myocardial infarction	*261 (3.3)*
In-hospital heart failure	*254 (3.2)*
Antiplatelet therapy at discharge	*416 (5.2)*
Statins at discharge	*416 (5.2)*
2-year prognostic model	
Age	*106 (1.7)*
Prior renal dysfunction	*260 (4.3)*
History of heart failure	*254 (4.2)*
AMI classification	*0 (0.0)*
Heart rate	*75 (1.2)*
Killip class	*54 (0.9)*
Hemoglobin	*135 (2.2)*
LVEF	*1161 (19.1)*
In-hospital PCI	*60 (1.0)*
In-hospital heart failure	*83 (1.4)*
Heart failure worsening within 30 days	*9 (0.1)*

Antiplatelet therapy within 30 days	*0 (0.0)*
β blockers within 30 days	*0 (0.0)*
Statins within 30 days	*0 (0.0)*

LVEF, left ventricular ejection fraction; PCI, percutaneous coronary intervention; AMI, acute myocardial infarction."

11) **Why have the authors not commented on the very low rate of COPD?** The rate of COPD is surprisingly low in this cohort (<1.9%) despite current smoker rate of c45% but no comment on this from the authors. Cf UK population of AMI the incidence of COPD is 12% (doi.org/10.1136/heartjnl-2016-309860). The incidence of prior heart failure is also surprisingly low (only c2%)

Response: Thank you for raising this interesting question. The rates of COPD and prior heart failure in our cohort were lower than the UK population of AMI. However, the rates in the present study were similar with data from the Improving Care for Cardiovascular Disease in China-Acute Coronary Syndrome (CCC-ACS) project, which was also a nationwide registry for ACS in China. The rates of COPD and prior heart failure were 1.3% (523/39915) and 0.9% (386/44563) in previous studies from CCC-ACS registry (doi: 10.1155/2021/9977312, doi: 10.1093/ehjacc/zuab053). The distinct prevalence of comorbidity in patients with myocardial infarction between countries highlighted the importance of developing, validating, improving, and applying different risk prediction models for different populations. We further added relevant discussion in the manuscript (**Page 17, Line 17-23**), as below.

"The rates of COPD and prior heart failure in our cohort were lower than the United Kingdom population of AMI.[30] However, the rates in the present study were similar with data from the Improving Care for Cardiovascular Disease in China-Acute Coronary Syndrome project, which was also a nationwide registry in China.[31,32] The distinct prevalence of comorbidity in patients with myocardial infarction between countries highlighted the importance of developing risk prediction model for specific population."

12) **Time dependent co-variates?** Is it truly a dynamic model – the authors prognosticate for 30 day and 2 year survival after AMI. They do not dynamically prognosticate in so far as one can only interpret probability of survival to two specific timepoints. In a dynamic model one would expect to be able to ascertain prognosis at any timepoint after the event for instance what about at 1 year after AMI? The model could be strengthen by including the variable "time since AMI" as a predictor. It is impossible to get away from survivor bias but in this prognostic model that is relevant – those who survive to 1 year after AMI are not the same as those who have only survived 30 days after AMI but they are both

represented the same in the 2 year prognostic model.

Response: We are grateful for this insightful comment. We totally agree that an optimal dynamic risk prediction model should be able to predict prognosis at any time points. However, from two aspects, we thought that the present models, which enabled risk assessment at discharge and risk reassessment at 30 days after discharge, also conformed to the conception of "dynamic" risk prediction to some extent. First, previous studies have found that a larger proportion of adverse events occurred in the early phase than at the late stage after AMI hospitalization (doi: 10.1136/heartjnl-2020-317165, doi: 10.1161/JAHA.120.019270), which was in line with observations in routine clinical practice and was also revealed in the present analysis. This means that it is more clinically meaningful to perform risk reassessment at an early time point after discharge from AMI (e.g., 30-day post-discharge) compared with any time points. Second, the present models "dynamically" integrated the clinical follow-up data of patients after discharge, making the risk prediction process to be longitudinal and dynamic. The existing "dynamic" models in patients with myocardial infarction, for example, the dynamic TIMI score (Dynamic TIMI risk score for STEMI. doi: 10.1161/JAHA.112.003269), only used variables obtained during hospitalization to achieve "dynamic" prediction of 1-year mortality at discharge. We thought with caution that the present risk prediction tool had already moved closer to the ideal dynamic risk prediction model for patients following AMI compared with previous work. We value this comment, and accordingly further revised the limitation section (**Page 19, Line 2-6**), as below.

"Second, although the present study showed the feasibility of assessing 2-year prognosis at 30 days after discharge, risk reassessment is a serial process and ideally performed at more time points beyond the early phase after discharge. Models which can ensure more dynamic and accurate risk prediction are still needed."

13) **Table 1.**

A minor point but inferential statistics such as a p-value in Table 1 is of limited to no use given the authors have already stated that they randomly allocated patients to their derivation or validation cohorts. There should be no reason to expect imbalanced groups unless failure of randomization. This is also true for Supplemental Table 1.

Age when presented in years should be a whole number unless the authors recorded the age in greater detail than whole years.

Heart rate should be reported in whole beats per minute.

Response: Thank you for these helpful suggestions for Table 1/online supplemental table 2 (1→2). Age was not presented in whole years because it was calculated as (admission time - date of birth)/365.25. We have addressed other problems in the tables (**Table 1; online supplemental table 2**).

14) **Spelling and Grammar – minor amendments.**
On line 22 page 12 it should read “than **GRACE** score” not “CRACE”

On line 21 page 18 it should read “informed **consent was** given” not “consents were”

The titles for Supplemental Figures 16 and 17 should begin:

“Comparison of Clinical **Utility**...”, the plural of utility in this context is still utility. Same for the use of the word utility in the caption text for Supplemental Fig 16 & 17.

The same is true of Line 9 page 12 it should read “better clinical **utility**”

Response: Thank you very much for your careful review. We have double checked the manuscript and corrected these errors.

15) **Regarding Tripod checklist.**
Item 1: Title does not state the risk being predicted: Mortality

Response: We have revised the title (**Page 1, Line 1-3**), as below.

“Development and validation of dynamic models to predict post-discharge mortality risk in patients with acute myocardial infarction: results from China Acute Myocardial Infarction registry”

Item 1: Title does not state that the population is patients who had already survived to hospital discharge

Response: We have revised the title (**Page 1, Line 1-3**), as below.

"Development and validation of dynamic prognostic models to predict post-discharge mortality risk in patients with acute myocardial infarction: results from China Acute Myocardial Infarction registry"

Item 8: Authors state the sample size but have not performed a sample size calculation. i.e. no justification obtaining sufficient events per predictor used in their model. A crude calculation from their 16 predictors (accounting for the multiple levels within their variables such as Killip classes) and 190 events in their 30 day model there are 11.8 events per predictor). It would be good practice for authors to show their sample size calculation/justify their sample size rather than just use all the data available.

Response: Thank you for this helpful suggestion. We further added our consideration of sample size in the method section (**Page 8, Line 16-22**), as below.

"In the derivation cohort, 190 deaths occurred within 30 days after discharge, which could ensure at most 19 predictor parameters (greater than 12 predictor parameters finally included) in the 30-day risk prediction model based on the rule of thumb that 10 events per candidate predictor parameters (EPP). Similarly, 740 deaths occurred between 30 days and 2 years, which could ensure at most 74 predictor parameters (greater than 15 predictor parameters finally included) in the 2-year risk prediction model.[15]"

Item 13: significant concerns about lack of missing data reporting. How it was handled (Item 9) is reported (multiple imputation) however the reason which in the case of missing prescribing data cases were excluded instead of imputed is not justified.

Response: The subgroup analyses were performed in the validation population with complete data on all model predictors instead of in the total study population, which was further clarified in the method section (**Page 10, Line 11-13**). The numbers of missing values for predictors in the risk prediction models were further summarized in **online supplemental table 1**. Only LVEF had >10% missing values in the total study population.

Item 18: Missingness is not discussed in the limitations particularly with respect to the variables used as predictors

Response: We further added the relevant content in the limitation section (**Page 19, Line 11 and 12**), as below.

"Fourth, there existed some missing values which needed to be imputed before regression analysis. However, almost all predictors had missing values of <6%.

VERSION 2 – REVIEW

REVIEWER	Luney, Matthew University of Nottingham, Anaesthesia & Critical Care
REVIEW RETURNED	09-Mar-2023

GENERAL COMMENTS	I would like to thank the authors for their considered and well written revised manuscript. They have comprehensively addressed the points raised in my review of their original submission. Reviewing their revised manuscript, in particular they have satisfactorily accounted for more recent models than GRACE 1.0, updated their limitations and discussion sections, and corrected the TRIPOD checklist responses. I am grateful for their clear additions to the supplement that detail missingness, and proportionality assessments. The discussion regarding hospital level findings (comment 9) is a particularly interesting finding that adds to the value of this manuscript in so far as it shows that this variable was not of prognostic value in the models presented. In conclusion I recommend to the editors to accept this manuscript for publication.

VERSION 2 – AUTHOR RESPONSE