PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Contemporary sex differences in mortality among patients with ST-segment elevation myocardial infarction: a systematic review and meta-analysis
AUTHORS	Qiu, Hong ; Xi, Ziwei; Guo, Tingting; Wang, Yong; Li, Jianan; Li, Yang; Zheng, Jianfeng; Gao, R

VERSION 1 – REVIEW

REVIEWER	Tereshchenko, Larisa Cleveland Clinic, Lerner Research Institute, Quantitative Health Sceinces
REVIEW RETURNED	27-Jul-2021

GENERAL COMMENTS	Xi et al. conducted a study with the objective to assess the effect of sex differences on short- and long-term mortality in pts with STEMI “by performing a meta-analysis.” the authors concluded that “ increased short-term mortality was found in women with STEMI. There are several major limitations in the study: inclusion criteria for studies were very wide, which resulted in many biases: except for one included study, heterogeneity by sex – analysis was not pre-specified, and thus statistical power was uncertain/low. Both RCTs and observational studies were included. There was a wide range of covariates included in adjusted models in various studies. The quality of the evidence reported by the included study was not accessed/not reported. The following questions need to be addressed: -abstract does not have appropriate sections per the journal style and has to be revised. -points #2 and #3 in Strength and Limitations have to be removed as they do not sound scientific as they point to expected study characteristics. -In Methods – please clarify which type of studies are included – only observational studies or RCTs? - what does it mean – “adjusted RRs described in included studies”? This is an incomplete description. Studies likely adjusted for a variety of covariates, and this is important. Please provide information regarding covariates/confounders that were included in the models for adjustment. How similar/different were they across the studies? In Results – - There was a wide range of covariates included in adjusted models in various studies. One potential solution could be to include in meta-analysis unadjusted point estimates, and then adjust meta-regression for the percentage of participants with confounders (hypertension, diabetes, etc, and average continuous

	variables (e.g. age)). Please consider such an approach as it would produce more reliable and less biased results. - there was no assessment of the quality of included studies. It is expected that the quality of included studies is assessed by the investigators - there were no assessment and clear reporting of missing data in the included studies. - the authors did not provide information about their study protocol, sources of funding, and competing interests. The “Other Information” subsection in the PRISMA checklist is incomplete.

REVIEWER	Bernal, José Hospital Universitario 12 de Octubre, Management Control
REVIEW RETURNED	03-Oct-2021

GENERAL COMMENTS	he authors conducted a systematic review and meta-analysis to evaluate the effect of sex differences on short- and long-term mortality among patients with ST-segment elevation myocardial infarction (STEMI). They conclude women with STEMI have an increased risk of short-term, but not long-term mortality, and the effect of sex differences remains significant after adjusting for cardiovascular risk factors and baseline clinical profiles. In all, the manuscript is compact and clear. It raises an important issue considering sex differences in mortality in STEMI patients, which is a relevant reference for future policies and interventions. However, in my opinion, some issues need to be clarified as they might affect the robustness of the results. I have the following queries and suggestions: Abstract 1. In the abstract, the authors state: "Risk ratio (RR) 95% CIs were measured using the Mantel-Haenszel method", but this method is not cited in the manuscript. There seems to be a discrepancy between the abstract and the manuscript that should be explained and corrected if necessary. 2. There also appears to be a mistake in the transcription of the results to the abstract in: "And adjusted long-term mortality was also similar between women and men (RR, 0.11; 95%CI, 0.42-1.80, P=0.008, I2=74.5%)", which the manuscript appears as: “(RR, 1.11; 95%CI, 0.42-1.80, P=0.008, I2=74.5%)”. 3. I suggest including the following aspects in the abstract, as recommended by PRISMA: background; more details on study eligibility criteria, study evaluation and synthesis methods; limitations and implications of the main findings. Methods 4. It seems clear the risk of 30-day mortality is greater than the risk of in-hospital mortality. To group both into a single category (short-term mortality) seems troublesome and I think requires an explanation. 5. I suggest a more detailed description of the inclusion/exclusion criteria of the studies, specifying whether their design (i.e., prospective vs. retrospective) data sources (i.e., clinical registry vs. administrative databases) and age ranges were considered. 6. It would also be interesting to group the results of the quality of the studies into categories according to the NOS scale, i.e., “the quality of the studies was divided into the following 3 categories:

	high quality (scores 7-9), moderate quality (scores 4-6) and low quality (scores 0-3)". Results 7. It is possible that, according to the authors' answers to questions 5 and 6, the results of the exclusions may require a more detailed description. 8. In my opinion, the authors should include in Table 1 the type of study, the data source and, if applicable, the variables considered for risk adjustment. 9. The results should be expressed systematically as numbers and percentages. 10. There is a mistake in: "We found no evidence of publication bias across studies based on visual inspection of funnel plots (See eFigure 3 in the Supplementary Material) and the results from Egger's tests for short-term mortality (P=0.462) and for long-term mortality (P=0.053)"? The authors state in the methods section: "To assess the potential effect of publication bias, we inspected funnel plots for asymmetry and used the Egger's regression asymmetry test in which P<0.10 was considered to indicate significant publication bias". 11. Some studies included in the meta-analysis use HR (Venetsanos et al.) and OR (i.e. Stehli et al and Tizon et al.) as measures of association. The authors express in the adjusted analysis RR, reflecting as RR values and their 95% confidence intervals exactly the same values that appear in the original studies expressed as HR or OR. I believe this treatment requires a justification. 12. Some studies (i.e., Stheli et al.) report mortality data for patients who have undergone PCI, but others (i.e., Hao et al.) report mortality data for STEMI patients in general, independently of whether they have undergone PCI. In my opinion, these are different populations whose comparison may be biased. The authors should justify this treatment, assess its implications and consider, in the event that they compromise the validity of their results, to what extent they should modify the study design. Discussion 13. In general, the authors do not address in the discussion questions related to the results of their study, but rather to the possible causes of the sex differences that they have found in their results, but which they have not analysed. In my opinion, the discussion should be reconsidered entirely. 14. In any case, I believe it would be appropriate to include in the discussion a comment regarding the interpretation of the result of the sensitivity analysis.

VERSION 1 – AUTHOR RESPONSE

We noted that reviewer #1 has commented on the strength and limitations section. Please note that this section should contain up to five short bullet points, no longer than one sentence each, that relate specifically to the methods of the study reported:
(see: http://bmjopen.bmj.com/site/about/guidelines.xhtml#articletypes).

It should not be a summary of the study. Thus, you might wish to revise the 2nd and 3rd bullet points of the strengths and limitations section after the abstract to explain why these are methodological strengths or limitations.

Response: Thank you for the comment. We have re-written the 2nd and 3rd bullet points of the strengths and limitations section in the revised manuscript.

Editorial requests:

- Please include the dates of the search in the abstract.

Response: We are very sorry for the lack of information on the dates of the search in the abstract. We added the dates of the search into the abstract in the revised manuscript as required.

Other changes: We have also revised the format of the Title page and corrected the affiliation of some co-authors (TT Guo and JN Li).

Responds to the Reviewer #1's comments:

-Abstract does not have appropriate sections per the journal style and has to be revised.

Response: Thank you for the comment. We are sorry for the incorrect structure of the abstract. We revised the abstract as required.

-points #2 and #3 in Strength and Limitations have to be removed as they do not sound scientific as they point to expected study characteristics.

Response: Thank you for this suggestion. We have re-written the 2nd and 3rd bullet points of the strengths and limitations section in the revised manuscript.

-In Methods

- please clarify which type of studies are included – only observational studies or RCTs?

Response: We are sorry for the unclear description about the inclusion criteria of relevant studies. Observational studies and RCTs were both included. After a comprehensive search and a careful screen for relevant studies, fifteen studies comprising of fourteen observational studies, except one post hoc analysis of randomized controlled trial, were included. The lack of eligible RCTs that that met the inclusion criteria was a major limitation in our analysis. We added more details on the inclusion criteria into the Methods section and the design of studies into the Table 1.

- what does it mean – "adjusted RRs described in included studies"? This is an incomplete description. Studies likely adjusted for a variety of covariates, and this is important. Please provide information regarding covariates/confounders that were included in the models for adjustment. How similar/different were they across the studies?

Response: We are very sorry for our incorrect writing and we made correction ("adjusted RRs described in included studies" was revised into "adjusted RRs if they were described in those included

studies"). Only 10 studies provided adjusted RRs and 9 studies reported what covariates/confounders their multivariate analyses adjusted for. Most studied adjusted for age, diabetes mellitus, hypertension, and prior myocardial infarction/PCI, while some adjusted for renal insufficiency, cardiogenic shock, cardiac arrest at admission and occurrence time of symptom onset (the 2nd paragraph of Results). As suggested by the Reviewer, we listed the variables adjusted in the adjusted analyses from the included studies in Supplementary Material (eTable 1). It needed to be noted that relevant confounders might differ greatly across studies. We added this into the Limitation part (the 8th paragraph of Discussion).

In Results –

- There was a wide range of covariates included in adjusted models in various studies. One potential solution could be to include in meta-analysis unadjusted point estimates, and then adjust meta-regression for the percentage of participants with confounders (hypertension, diabetes, etc, and average continuous variables (e.g. age)). Please consider such an approach as it would produce more reliable and less biased results.

Response: Many thanks for this comment. It is really true as the Reviewer suggested that meta-regression was a well-accepted approach to investigate the sources of heterogeneity in the included studies. We performed meta-regression for available confounders provided by included studies and added the results of meta-regression into the Results in the revised manuscript as the Reviewer suggested. However, considering that part of included studies did not provide information on confounders stratified by sex, it should be noted that meta-regression might not be appropriate when less than 10 studies were incorporated into calculation.

- there was no assessment of the quality of included studies. It is expected that the quality of included studies is assessed by the investigators

Response: Thank you for the comment. The quality of included studies was assessed by Newcastle-Ottawa scale and the results of assessment was presented in the eTable 2 of the Supplementary Material. In addition, we added stratified analysis according to dividing the included studies into different subgroups based on the NOS scores (>7 points or ≤7 points) (the 7th paragraph of the Methods section and the 4th paragraph of the Results section).

- there were no assessment and clear reporting of missing data in the included studies.

Response: We added description on missing data of the included studies into the Results (the 2nd paragraph of the Results) as suggested. All included studies provided sufficient data for analysis of sex differences in clinical outcomes, while baseline characteristics of participants in some of them were lacking.

- the authors did not provide information about their study protocol, sources of funding, and competing interests. The "Other Information" subsection in the PRISMA checklist is incomplete.

Response: Thank you for the reminder. We added the section of "Other Information" following Discussion section in the revised manuscript. This research received no funding agency. The authors

declare that there is no conflict of interest. The institutional review board central committee of our center approved that the study protocol and inform consent was not required for our study.

Special thanks to you for your good comments.

Responds to the Reviewer #2's comments:

Abstract

1. In the abstract, the authors state: "Risk ratio (RR) 95% CIs were measured using the Mantel-Haenszel method", but this method is not cited in the manuscript. There seems to be a discrepancy between the abstract and the manuscript that should be explained and corrected if necessary.

Response: We are very sorry for our negligence this incorrect writing. The data were combined using the DerSimonian and Laird random-effects model in our study. In the beginning, we thought that a fixed-effects model using the Mantel–Haenszel's method might be appropriate. However, wo then found that there was substantial heterogeneity in our analysis. Thus, we finally chose the DerSimonian and Laird random-effects model. We made correction in the abstract in the revised manuscript.

2. There also appears to be a mistake in the transcription of the results to the abstract in: "And adjusted long-term mortality was also similar between women and men (RR, 0.11; 95%CI, 0.42-1.80, P=0.008, I2=74.5%)", which the manuscript appears as: "(RR, 1.11; 95%CI, 0.42-1.80, P=0.008, I2=74.5%)".

Response: We apologize for the mistake in the original manuscript. We have made correction in the revised manuscript.

3. I suggest including the following aspects in the abstract, as recommended by PRISMA: background; more details on study eligibility criteria, study evaluation and synthesis methods; limitations and implications of the main findings.

Response: Thank you for this suggestion. To be in accordance with the journal style and as the Reviewer suggested, we have added more details into the abstract and revised the format of this part.

Methods

4. It seems clear the risk of 30-day mortality is greater than the risk of in-hospital mortality. To group both into a single category (short-term mortality) seems troublesome and I think requires an explanation.

Response: It is really true as the Reviewer pointed out that the risk of 30-day mortality is definitely higher than the risk of in-hospital mortality and putting them into one category seems not so appropriate. However, in some of the included studies (i.e., Hannan et al. and Stehli et al.), both 30-day and in-hospital mortality were directly classified into short-term mortality and they did not provide sufficient information to calculate the 30-day and in-hospital mortality respectively. Thus, both of them were consistently classified into short-term mortality in our analysis.

5. I suggest a more detailed description of the inclusion/exclusion criteria of the studies, specifying whether their design (i.e., prospective vs. retrospective) data sources (i.e., clinical registry vs.

administrative databases) and age ranges were considered.

Response: We deeply appreciate the reviewer's suggestion. To be more clearly and in accordance with the reviewer concerns, we added more details on the inclusion/exclusion criteria into the Methods section. Both prospective and retrospective studies, and studies of clinical registry and administrative databases were included in our analysis. There was no age limit in our analysis.

6. It would also be interesting to group the results of the quality of the studies into categories according to the NOS scale, i.e., "the quality of the studies was divided into the following 3 categories: high quality (scores 7-9), moderate quality (scores 4-6) and low quality (scores 0-3)".

Response: We agree with the Reviewer and we added stratified analysis according to dividing the included studies into different subgroups based on the NOS scores (>7 points or ≤7 points) (the 7th paragraph of the Methods section and the 4th paragraph of the Results section). In the subgroup analysis, the RR unadjusted short-term mortality of studies with NOS >7 points were numerically higher than that of studies with≤7 points (See eFigure 1 in the Supplementary Material). Given that only one study has a Newcastle-Ottawa scale >7 points among studies reporting long-term all-cause mortality, we did not perform subgroup analysis according to the NOS scores for long-term mortality.

Results

7. It is possible that, according to the authors' answers to questions 5 and 6, the results of the exclusions may require a more detailed description.

Response: We are grateful for above suggestion and have added more details on the inclusion/exclusion process of the studies into the Methods and Results sections. We also added the results of stratified analysis based on the quality of the included studies into the Results section.

8. In my opinion, the authors should include in Table 1 the type of study, the data source and, if applicable, the variables considered for risk adjustment.

Response: According to this comment, we added the type of study and the data source into the Table 1. We also listed the variables adjusted in the adjusted analyses from the included studies in Supplementary Material (eTable 1) in the revised manuscript as suggested.

9. The results should be expressed systematically as numbers and percentages.

Response: Thank you for the comment. We have expressed the results as numbers and percentages including the baseline characteristics in Table 2 in the revised manuscript as suggested.

10. There is a mistake in: "We found no evidence of publication bias across studies based on visual inspection of funnel plots (See eFigure 3 in the Supplementary Material) and the results from Egger's tests for short-term mortality (P=0.462) and for long-term mortality (P=0.053)"? The authors state in the methods section: "To assess the potential effect of publication bias, we inspected funnel plots for asymmetry and used the Egger's regression asymmetry test in which P<0.10 was considered to indicate significant publication bias".

Response: We are very sorry for this incorrect writing in the original manuscript. P< .05 was considered to be statistically significant for the Egger's regression asymmetry test (refer to: Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629). We have made the correction in the revised manuscript.

11. Some studies included in the meta-analysis use HR (Venetsanos et al.) and OR (i.e. Stehli et al and Tizon et al.) as measures of association. The authors express in the adjusted analysis RR, reflecting as RR values and their 95% confidence intervals exactly the same values that appear in the original studies expressed as HR or OR. I believe this treatment requires a justification.

Response: Thank you for this comment. Just as Reviewer pointed out, some studies used HRs and ORs, which was inconsistent with RRs used in our meta-analysis{RR = [A/(A+B)] ÷[C/(C+D)] = A(C+D) /C(A+B), OR = (A/B)÷(C/D) = AD/BC}. However, if incidence of events (A and C) was relatively low (usually < 20%), ORs could be directly used as RRs (OR≈RR). Additionally, in meta-analysis, HRs could be used interchangeably with RRs. Thus, it was reasonable to express the results in RRs in our study while the included studies used HRs or ORs.

12. Some studies (i.e., Stheli et al.) report mortality data for patients who have undergone PCI, but others (i.e., Hao et al.) report mortality data for STEMI patients in general, independently of whether they have undergone PCI. In my opinion, these are different populations whose comparison may be biased. The authors should justify this treatment, assess its implications and consider, in the event that they compromise the validity of their results, to what extent they should modify the study design.

Response: We agree with the Reviewer that patients treated with reperfusion therapy and no-reperfusion therapy were different population and might have various prognosis. But a meta-analysis from Pancholy et al. (JAMA Intern Med. 2014;174(11):1822-1830.) has examined differences in mortality by sex in patients with STEMI treated with primary PCI. Its results demonstrated that women were at a higher risk for in-hospital and 1-year all-cause mortality compared with men and the higher risk for 1-year mortality in women was no longer significant when adjusted RRs were used, which was completely consistent with our study. The increasing use of primary PCI and the growing rate of reperfusion therapy in recent years could partly explained the consistency of our study and previous studies conducted among patients undergoing PCI. We added some comments about on this into the Discussion (the 6th paragraph of the Discussion section). Besides, impact of contemporary invasive versus conservative treatment strategies in women and men with STEMI is the topic of our future studies.

Discussion

13. In general, the authors do not address in the discussion questions related to the results of their study, but rather to the possible causes of the sex differences that they have found in their results, but which they have not analyzed. In my opinion, the discussion should be reconsidered entirely.

Response: Thank you for the comment. We added more discussion on our results (the 2nd and 6th paragraph of Discussion section) as suggested.

14. In any case, I believe it would be appropriate to include in the discussion a comment regarding the interpretation of the result of the sensitivity analysis.

Response: Thank you for this suggestion. Our sensitivity analysis suggested consistent results of short-term mortality. Considering that the sensitivity analysis showed a significant association between female and increased long-term mortality after removing one study from adjusted analyses, we believed that the finding of non-significant increased long-term mortality in adjusted analyses of our study need be interrupted with caution. We added more discussion on sensitivity analysis as suggested (the 2nd paragraph of Discussion).

VERSION 2 – REVIEW

REVIEWER	Tereshchenko, Larisa Cleveland Clinic, Lerner Research Institute, Quantitative Health Sceinces
REVIEW RETURNED	13-Nov-2021

GENERAL COMMENTS	The manuscript significantly improved. The authors adequately addressed all questions. One point remains: On page 15, there is an error: it is stated that RR 1.5 (1.23-1.83), which implies a statistically significant finding. However, P-value is reported as 0.148 and is non-significant. Something is wrong here.

REVIEWER	Bernal, José Hospital Universitario 12 de Octubre, Management Control
REVIEW RETURNED	23-Nov-2021

GENERAL COMMENTS	1. If, as the authors admit, the risk of 30-day mortality is decidedly greater than the risk of in-hospital mortality and putting them in a single category does not seem so appropriate, the reason for doing something inappropriate by putting both groups in a single category can never be that some of the studies considered do not differentiate the two outcome variables. To avoid bias, I understand that it would be advisable: 1. To not include studies that do not provide sufficient information or 2: if not, to group the outcome variables together. 2. I do not quite understand the solution provided to the question of adjusted RRs. I think it is not only a matter of clarifying that they have been calculated using adjusted RRs if they were described in those included studies, but also, as proposed by reviewer 1, alternatively including unadjusted point estimates in the meta-analysis, and then adjusting the meta-regression for the percentage of participants with confounding factors.

	In any case, it seems to me that the meta-regression performed should be better detailed and parameters describing the heterogeneity of the model, such as τ2, τ, I2, H2, R2adj or Q, should be included in the manuscript. 3. If patients treated with reperfusion therapy and without reperfusion are different populations, the fact that the results of the meta-analysis by Pancholy et al. coincide with the authors' results does not seem to me to resolve the question posed, which remains whether both populations should be treated as they were the same.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

The manuscript significantly improved. The authors adequately addressed all questions.

One point remains:

On page 15, there is an error: it is stated that RR 1.5 (1.23-1.83), which implies a statistically significant finding. However, P-value is reported as 0.148 and is non-significant. Something is wrong here.

Response: We are very sorry for our negligence. The P value of the RR is <0.001 and we provided the P value of heterogeneity chi-squared by mistake. We made the correction in the revised manuscript. Special thanks to you for your comments, which are very valuable and helpful.

Reviewer 2

1. If, as the authors admit, the risk of 30-day mortality is decidedly greater than the risk of in-hospital mortality and putting them in a single category does not seem so appropriate, the reason for doing something inappropriate by putting both groups in a single category can never be that some of the studies considered do not differentiate the two outcome variables. To avoid bias, I understand that it would be advisable: 1. To not include studies that do not provide sufficient information or 2: if not, to group the outcome variables together.

Response: Thanks for the comment. To be more clearly and in accordance with the Reviewer's concerns, we divided the short-term mortality into 30-day and in-hospital mortality, and calculated the RR for 30-day and in-hospital mortality, respectively (the 6th paragraph of the Methods section and the 5th paragraph of the Results section). Among 13 studies which investigated the sex differences on short-term outcomes after STEMI, 7 studies reported sex-specific mortality at 30 days after SETMI and 6 studies provided data on the in-hospital mortality. We found that the impact of sex on in-hospital and 30-day mortality were consistent.

2. I do not quite understand the solution provided to the question of adjusted RRs. I think it is not only a matter of clarifying that they have been calculated using adjusted RRs if they were described in those included studies, but also, as proposed by reviewer 1, alternatively including unadjusted point

estimates in the meta-analysis, and then adjusting the meta-regression for the percentage of participants with confounding factors.

In any case, it seems to me that the meta-regression performed should be better detailed and parameters describing the heterogeneity of the model, such as τ2, τ, I2, H2, R2adj or Q, should be included in the manuscript.

Response: We are grateful for the suggestion. As suggested by the Reviewer, we have added more details on the results of the meta-regression (the 7th paragraph of the Results section).

3. If patients treated with reperfusion therapy and without reperfusion are different populations, the fact that the results of the meta-analysis by Pancholy et al. coincide with the authors' results does not seem to me to resolve the question posed, which remains whether both populations should be treated as they were the same.

Response: Many thanks for this comment. Considering the Reviewer's concern on whether patients treated with and without reperfusion were different populations, we performed an additional meta-analysis of studies (short-term: Ali et al., Stehli et al and Kerkman et al; long-term: Tang et al. Kerkman et al.) which included patients undergoing PCI (short-term: RR 1.45, 95%CI 1.05-2.00, P=0.026, I2=39.5%; long-term: RR 1.28, 95%CI 0.95-1.73, P=0.108, I2=0.0%) (the 5th and 6th paragraph of the Results section). Due to insufficient data on patients treated with no reperfusion therapy, the comparison between the two populations was not available in our study. In addition, even though there was difference in the outcomes of patients treated with reperfusion therapy and without reperfusion, we should consider patients with STEMI as a whole which contained patients receiving procedure or not.

We are really grateful to the Reviewer for his effort reviewing our paper.

VERSION 3 – REVIEW

REVIEWER	Tereshchenko, Larisa Cleveland Clinic, Lerner Research Institute, Quantitative Health Sceinces
REVIEW RETURNED	30-Nov-2021

GENERAL COMMENTS	no additional questions

REVIEWER	Bernal, José Hospital Universitario 12 de Octubre, Management Control
REVIEW RETURNED	30-Nov-2021

GENERAL COMMENTS	I appreciated the responses and I don't have further comments.