PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prognostic Significance of Platelet-Lymphocyte Ratio (PLR) in Breast Cancer Patients Treated with Neoadjuvant Chemotherapy: A Meta-analysis
AUTHORS	Qi, Xue; Chen, Jia; Wei, Sheng; Ni, Jingyi; Song, Li; Jin, Conghui; Yang, Lei; Zhang, Xunlei

VERSION 1 – REVIEW

REVIEWER	Yamanouchi, Kosho National Hospital Organisation Nagasaki Medical Center, Surgery
REVIEW RETURNED	09-Jun-2023

GENERAL COMMENTS	The evaluation was significant. The results, description of needs for further study, and limitation of meta analysis were well written notably.

REVIEWER	Gianni, Caterina IRCCS Istituto Romagnolo per lo Studio dei Tumori Dino Amadori
REVIEW RETURNED	12-Jun-2023

GENERAL COMMENTS	The article is well written and interesting. It completes the information available about the topic. Minor english revision is required. I think the autors should a extend the part of the text on limitations of the study in particular for the high heterogeneity of studies included in the analysis. Other comments are available in the file attached

REVIEWER	Gwon, Yeongjin University of Nebraska Medical Center, Biostatistics
REVIEW RETURNED	10-Aug-2023

GENERAL COMMENTS	This study is well-written, and the findings presentation is also well-described. I have modest reservations about conducting meta-analysis, particularly stratification analysis. However, this does not interfere with my understanding of the article's primary points. I believe the team should be able to quickly address this. To analyze heterogeneity among trials, the authors used the I2 statistic and Cochran's Q statistic, followed by the P-value, P_h. P_h is based on Q test statistics, which means that when the number of studies is large, the values are more accurate. When looking at the overall meta-analysis, it would be OK to use P_h to determine the significance of heterogeneity. However, in most situations, the stratification did not exceed five. I was wondering if

	the authors were able to suggest alternative methods for testing for heterogeneity. Lines 58-60 on page 9: Could the authors explain why they chose 0.10 as the p-value threshold? Also, I2 is a measure of heterogeneity, it should be stated as "substantial" rather than "significant." After running a statistical test, the term "significant" can be employed.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:
Comment 1: The evaluation was significant.
The results, description of needs for further study, and limitation of meta analysis were well written notably.
Answer: Thanks for your review and positive recommendation.

Reviewer #2:
Comment 1: The article is well written and interesting. It completes the information available about the topic.
Minor english revision is required.
I think the autors should a extend the part of the text on limitations of the study in particular for the high heterogeneity of studies included in the analysis.
Other comments are available in the file attached
Answer: Thanks for your carefully review about our manuscript. We have revised the entire manuscript carefully according to your suggestions marked in the article including minor english revision (see attached file: "Main Document - marked copy"). Besides, we extend the part of the text on limitations of the study as highlited in the Main Document.

Reviewer #3:
Comment 1: To analyze heterogeneity among trials, the authors used the I2 statistic and Cochran's Q statistic, followed by the P-value, P_h. P_h is based on Q test statistics, which means that when the number of studies is large, the values are more accurate. When looking at the overall meta-analysis, it would be OK to use P_h to determine the significance of heterogeneity. However, in most situations, the stratification did not exceed five. I was wondering if the authors were able to suggest alternative methods for testing for heterogeneity.
Answer: Thanks for your carefully review.
We know that the Q-test has certain limitations, and its results are sometimes unstable, especially in stratified analysis or when the number of literature is small. So we also used the I2 statistic in this article to analyze the heterogeneity, which can correct the impact of the number of literature on the Q statistic. In the stratified analysis of this article, the results of I2 statistic were more accurate than Q statistic. Moreover, in addition to the I2 and Q statistic, we can also use the H statistic, Galbraith plot method and L'Abe plot method to analyze heterogeneity.

Comment 2: Lines 58-60 on page 9: Could the authors explain why they chose 0.10 as the p-value threshold?
Answer: Thanks for your suggestions. In our opinions, the p-value threshold usually set as 0.05 or 0.10 were acceptable. As reasons metioned in the above question, We set 0.10 as the p-value threshold to expand its testing efficiency.

Comment 3: Also, I2 is a measure of heterogeneity, it should be stated as "substantial" rather than

"significant." After running a statistical test, the term "significant" can be employed.
Answer: Thanks for your suggestions. We have correct the description as followes: If the P value of the Q test was <0.10 or I2 >50%, indicating significant or substantial heterogeneity across studies......

VERSION 2 – REVIEW

REVIEWER	Gwon, Yeongjin University of Nebraska Medical Center, Biostatistics
REVIEW RETURNED	27-Oct-2023

GENERAL COMMENTS	I appreciate the authors' efforts to respond to the feedback and comments of the reviewers. Unfortunately, the authors' selection of distinct p-value thresholds for heterogeneity and effect magnitude continues to puzzle me. "We set 0.10 as the p-value threshold to expand its testing efficiency," the authors state in their response. This, however, is not true to the best of my knowledge. In the decision-making, it is essential that the threshold for all statistical tests be consistent. Furthermore, expanding the tolerance for type I errors from 0.05 to 0.1 does not influence the efficacy of the statistical test and will inevitably result in erroneous conclusions. It is highly recommended that the authors take into account consistent thresholds. I would not be surprised if this had an impact on some of their findings, but that is perfectly acceptable.

VERSION 2 – AUTHOR RESPONSE

Reviewer #3:
Comment 1: I appreciate the authors' efforts to respond to the feedback and comments of the reviewers. Unfortunately, the authors' selection of distinct p-value thresholds for heterogeneity and effect magnitude continues to puzzle me. "We set 0.10 as the p-value threshold to expand its testing efficiency," the authors state in their response. This, however, is not true to the best of my knowledge. In the decision-making, it is essential that the threshold for all statistical tests be consistent. Furthermore, expanding the tolerance for type I errors from 0.05 to 0.1 does not influence the efficacy of the statistical test and will inevitably result in erroneous conclusions. It is highly recommended that the authors take into account consistent thresholds. I would not be surprised if this had an impact on some of their findings, but that is perfectly acceptable.
Answer: Thanks for your suggestion for the selection of distinct p-value thresholds for heterogeneity. After seriously discussion, we think that your opinion is correct and helpful for us. We realized that setting 0.10 as the p-value threshold for heterogeneity may result in the inconsistance of statistical tests and expanding the tolerance for type I errors. So we correct the state as follows: If the P value of the Q test was <0.05 or I2 >50%, indicating significant heterogeneity across studies......Fortunately, after re-examining our results, we found that this change did not affect our main conclusion. The changes were highlighted in the file "Main Document - marked copy".

VERSION 3 – REVIEW

REVIEWER	Gwon, Yeongjin University of Nebraska Medical Center, Biostatistics
REVIEW RETURNED	07-Nov-2023

| **GENERAL COMMENTS** | No further comment. |