## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Indirect Comparison between Immunotherapy Alone and Immunotherapy plus Chemotherapy as First-Line Treatment for Advanced Non-Small Cell Lung Cancer: A Systematic Review |
| **AUTHORS** | Li, Lingling; Xu, Fei; Chen, Yu; Ren, Xiaoli; Liu, Yu; Chen, Yuan; Xia, Shu |

## VERSION 1 – REVIEW

| | |
|---|---|
| **REVIEWER** | Martin Früh<br>Oncology and Hematology<br>Kantonsspital St. Gallen<br>9007 St. Gallen |
| **REVIEW RETURNED** | 30-Sep-2019 |

| | |
|---|---|
| **GENERAL COMMENTS** | This metaanalysis investigates the question of immunotherapy (IO) vs. immunotherapy plus chemotherapy (IC) in metastatic NSCLC. The study is well written and the methods appear accurate and well described.<br><br>I have however some important points to consider<br>1) The essential clinical question is whether to use IO vs IC in patients with metastastic NSCLC and PD-L1 expression levels of > or equal to 50%. According to Keynote 042, which was a large study, the general belief is, that in the subgroup PD-L1 levels 1-49% actually chemotherapy is preferrable, thus IC is likely also preferrable in this population over IO alone. I recommend to focus the analysis and the reporting of the results/ discussion primarly on the group of PD-L1 high. No one is giving first line IO alone to PD-L1 unselected patients and PD-L1 testing is a current standard.<br>2) The discussion is not really a discussion but rather a repetition of the results. I suggest to add discussion points such as the meaning/interpretation of the findings and its potential consequences for current practise/future studies (i.e. why are females/never smokers doing better with IC, also one could more elaborate about decreased irAEs in the ICI group which appears to be a new finding (biologic rational)<br>3) The results/interpretation of the abstract doesn t match the reporting in the results/discussion/conclusion in the manuscript. I suggest to rewrite the abstract accordingly |

| | |
|---|---|
| **REVIEWER** | Danilo Rocco<br>AORN dei Colli Monaldi, Division of Pulmonary Oncology<br>Naples. Italy |
| **REVIEW RETURNED** | 23-Jan-2020 |

| | |
|---|---|
| **GENERAL COMMENTS** | This article addresses a very relevant and timely topic (IO vs. IC in first-line advanced NSCLC treatment) and could surely benefit the current state of the art. The study objective of this paper is clearly defined, the abstract is accurate and complete. The study design is appropriate and the methods are described sufficiently. The outcomes and results are clearly addressed and the statistics are used and described appropriately; the references are up-to-date. The discussion and conclusions are justified by the results presented.<br>However, I woukd suggest some corrections:<br><br>1) This article reads poorly for an English audience and is difficult to follow, thus needs substantial grammar and language revision throughout the manuscript, especially from pages 8 to 15 and 15 to 18.<br>2) Every time you say "advanced NSCLC", please clarify to which subset of NSCLC you are referencing.<br>3) Page 2 lines 30-31: With reference to conference abstracts, I would add that you searched ClinicalTrials.gov, American Society of Clinical Oncology Meeting Library, and World Conference on Lung Cancer to collect them.<br>4) Page 4 lines 11-12: I would rephrase as 85% of lung cancer malignancies<br>5) Page 4 lines 17-18: I would specify which drivers<br>6) Page 4-5: KEYNOTE-024 and 189 need to be more comprehensively discussed, please add their design, number of patients, and outcomes (PFS and OS in months, TRAEs, ORR etc.)<br>7)I think the discussion section would benefit from the addition of the currently available metanalyses on this topic (e.g. Doherty 2019 et al., Zhou et al 2019), please add, compare and discuss<br>8) Figure 2: Please add the subgroup to better clarify what you are comparing |

| | |
|---|---|
| **REVIEWER** | Yiwei Zhang<br>Merck & Co., USA |
| **REVIEW RETURNED** | 06-Mar-2020 |

| | |
|---|---|
| **GENERAL COMMENTS** | The paper compares efficacy and safety of immunotherapy (IO) with immunotherapy plus chemotherapy (IC) for first-line advanced NSCLC via indirect comparison. My key comments are as follows:<br><br>1. The conclusion in the abstract is not very rigorous. It should be restricted to results by indirect comparison and certain efficacy and safety measures.<br>2. The cited paper about indirect comparison mentioned that “the validity of the adjusted indirect comparisons depends on the internal validity and similarity of the included trials”. The authors should justify the validity and similarity of the included trails, in particular the baseline characteristics of each trial.<br>3. The statistical analysis methods are not very clear. The authors don’t specify what specific meta-analysis method is used; more importantly, how the subgroup analysis is conducted. I don’t think all the trials reported the efficacy and safety results in each subgroup. Since the conclusion is mostly from the subgroup analysis results, the authors need to talk about the methods explicitly.<br>4. Some terminologies are not very accurate. For example, in page 10 under Overall Survival, the pooled analysis usually means |

| | |
|---|---|
| | to pool individual data together, but I don't think the authors have individual level data for each clinical trial.<br>5. Supplemental figures are not labeled correctly. |

| | |
|---|---|
| **REVIEWER** | Sangchoon Jeon<br>Yale University, United States |
| **REVIEW RETURNED** | 30-Apr-2020 |

| | |
|---|---|
| **GENERAL COMMENTS** | This study was performed in very formalized manor and well organized in order to examine the immnuno only over immune with chemo using indirect comparisons. However, there are few minor concerns.<br>1. The sample sizes are so different by the analysis. For Figure 3 (overall survival), 1,815 subjects are in immuno vs. chemo and 1,854 in immuno+chemo vs. chemo. They are pretty similar sameple sizes. However, for Figure 4 (progression-free survival), 846 subjects are in immuno only vs. chemo and 3,114 subjects are in immuno+chemo vs. chemo. The samples in immuno only decreased more than 50% while the samples in immuno+chemo increased 67%. For Figure 5 (adverse events), 2,085 and 3,579 subjects were included in immune only and immuno+chemo studies respectively. There is no explanation about this discrepancy. Need to clarify the sample size differences with rationale. Also we need to be sure if the discrepancy did not lead bias.<br>2. Authors performed to test heterogeneity across studies, but I recommend they need to clarify which variables or outcome measures they are compared. In addition, it would be good to know if the control groups (chemo only) between the studies to compare immuno only and immuno + chemo had similar results in terms of HR, RR, and adverse events. For indirect comparison, we need to be sure the control groups had similar performance. |

## VERSION 1 – AUTHOR RESPONSE

Responses to Reviewers
To Reviewer 1:
1. Considering the Reviewer's suggestion, we have re-written the results/discussion primarly on the group of PD-L1 high and just made a brief summary of PD-L1 at least 1-49% and at least 1%. We deleted the relevant passage according to the Reviewer's suggestion. (page10, page11 and page14).
2. The discussion was completely rewritten. We have added the following discussion to talk about why are females/never smokers doing better with IC, decreased irAEs in the ICI group (page 15).
3. We are very sorry for our negligence of the difference between results/interpretation in the abstract and results/discussion/conclusion in the manuscript. We have re-written the abstract to keep results consistent. (page2)
To Reviewer 2:
1. We regret there were problems with the English. The paper has been carefully revised by a professional language editing service (www.editage.cn) to improve the grammar and readability.
2. We have clarified the subset of the "advanced NSCLC" which refers to "stage IIIB and IV". (page 2 line 3)
3. Page 2 lines 8-10: We have added that you searched ClinicalTrials.gov, American Society of Clinical Oncology Meeting Library, and World Conference on Lung Cancer to collect them.
4. Page 4 lines 4: We have rephrased as 85% of lung cancer malignancies.

5. Page 4 lines 6-7: We have specified which drivers.
6. Page 4-5:We have discussed KEYNOTE-024 and 189 comprehensively and added their design, number of patients, and outcomes (PFS and OS in months, TRAEs, ORR etc.)
7. The meta-analysis of Doherty et al. 2019 and Zhou et al. 2019 have been added to the text and Reference section. (page 14-15)
8. We are extremely grateful to Reviewer 2 for pointing out this problem. We have revised Figure 2 to clarify that this picture is the comparison of summary Immunotherapy alone and Immunotherapy plus chemotherapy in OS, PFS and ORR including the results of subgroup analysis.
To Reviewer 3:
1. We have made correction about the conclusion in the abstract according to the Reviewer's comments. (page2)
2. We have justified the validity and similarity of the included trials including the baseline characteristics of each trial in the page17.
3. We have added the statistical analysis methods and the method of subgroup analysis in the page 9.
4. The pooled analysis (page10 line 8) was incorrectly stated in the original manuscript. This has been rectified. We are grateful to the referees for pointing out their error.
5. We have added the label of the legend of Supplemental figures correctly in page25.

To Reviewer 4:
1. As Reviewer suggested that, a discussion of the sample size differences with rationale has been included. This might lead to the imbalance of the patient population to affect the comparability of the indirect comparison and thus produce a potential selection bias (page17).
2. 2.1 In terms of variables or outcome measures that we are compared to test heterogeneity across studies, we have clarified that statistical heterogeneity in the included studies was evaluated using the chi-squared test and I2 statistic. When I2 was < 50% and p was > 0.1, a fixed effects model was selected to combine the studies; otherwise, a random effects model was used. Sensitivity analysis was conducted to explain the heterogeneity. (page8)
2.2 We have added the reason that the control groups have similar performance (page 16).
We tried our best to improve the manuscript and made some changes in the manuscript. These changes will not influence the content and framework of the paper. And here we did not list the changes but marked in red in revised paper.
We appreciate for Editors/Reviewers' warm work earnestly, and hope that the correction will meet with approval.
Once again, thank you very much for your comments and suggestions.

**VERSION 2 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Yiwei Zhang<br>Merck & Co., USA |
| **REVIEW RETURNED** | 30-Jun-2020 |

| | |
|---|---|
| **GENERAL COMMENTS** | I think this manuscript is still not complete. My key comments are as follows:<br><br>1. Based on the results, the benefits of IC vs IO in patients with PD-L1 expression at least 1% is more than the benefits of IC vs IO in patients with PD-L1 expression at least 50%. Why do the authors only mention about the subgroup of PD-L1 expression at least 50% in the conclusions? Is there an explanation why PD-L1 >=1% subgroup had better IC benefits over IO than PD-L1 >=50% subgroup? |

| | |
|---|---|
| | 2. The statistical analysis methods are still not very clear. Did all the trails included report the efficacy and safety results in each subgroup, such as PD-L1 >= 1%, PD-L1 1-49%, and PD-L1 >= 50%? If yes, why the authors do not report the data for all PD-L1 subgroups? For example, the PFS for PD-L1 1-49% subgroup is not reported. I assume the PD-L1 Low is defined as PD-L1 1-49%, but it is better to clarify it in the text. In addition, is there any data for PD-L1 < 1% subgroup? The results for PD-L1 subgroups are not very organized to me; whereas the main conclusion is drawn from here.<br>3. The figures 1-5 are missing in the main context. |

**VERSION 2 – AUTHOR RESPONSE**

Responses to Reviewer

To Reviewer 3:

1. We are extremely grateful to Reviewer 3 for pointing out this problem.

1.1 Based on the safety and superior survival outcomes reported in the phase III KEYNOTE-024 trial, the United States Food and Drug Administration approved pembrolizumab monotherapy as a first-line treatment for patients with advanced NSCLC whose tumor express a PD-L1 tumor proportion score (TPS) ≥50%. PD-L1 testing is a current standard. The essential clinical question is whether to use IO vs. IC in patients with metastatic NSCLC and PD-L1 expression in at least 50% of tumor cells. Therefore, this meta-analysis focused the analysis and the reporting of the results/ discussion primarily on the group of PD-L1 high expressed (≥50%).

1.2 There were fewer randomized controlled trials of first-line treatment for advanced NSCLC with current findings, which led to few studies being included in this analysis, especially for IO. In addition, not all studies reported the outcome indicators in this meta-analysis, and the sample sizes were different between IC and IO. This might lead to the imbalance of the patient population to affect the comparability of the indirect comparison and thus produce a potential selection bias. Furthermore, the data of PD-L1 ≥1% subgroup and PD-L1 ≥50% subgroup analyses were not from the same studies.

2. We have made correction about the statistical analysis according to the Reviewer's comments.

2.1 We have added the method of subgroup analysis in the page 9. According to PD-L1 expression, the main subgroup included PD-L1 high expressed subgroup (≥50%), PD-L1 low expressed subgroup (1%–49%), and PD-L1 positive subgroup (≥1%). Due to lack of data, the subgroup of PD-L1 expression less than 1% was not performed. Not all the trials reported the efficacy and safety results in each subgroup. We extracted the subgroup analysis data of all the trials according to the pre-designed grouping factors, and each trial was included only once per subgroup.

2.2 We have clarified the definition of PD-L1 high expressed subgroup (≥50%), PD-L1 low expressed subgroup (1%–49%), and PD-L1 positive subgroup (≥1%). (page9)

2.3 Due to lack of data, the subgroup of PD-L1 expression less than 1% was not performed. For PD-L1 negative subgroup (<1%), immunotherapy plus chemotherapy versus chemotherapy, for example, KEYNOTE-407 and KEYNOTE-189 had related statistical analysis. However, the clinical trials of immune single-agent versus chemotherapy for first-line therapy had not found an analysis of this part of the population. Therefore, there was no indirect comparison between immunotherapy alone and immunotherapy plus chemotherapy for PD-L1 <1%.

3. We re-uploaded Figure 1-5, and checked the result section including description and mention Figure 1-5.

We appreciate for Editors/Reviewers' warm work earnestly, and hope that the correction will meet with approval.

Once again, thank you very much for your comments and suggestions.

**VERSION 3 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Yiwei Zhang<br>Merck & Co., USA |
| **REVIEW RETURNED** | 21-Aug-2020 |

| | |
|---|---|
| **GENERAL COMMENTS** | Thank you for responding to my questions. I don't have additional comments. |