PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prospective, multicenter, single-arm phase II trial of pembrolizumab combined with carboplatin and pemetrexed in elderly patients with advanced, non-squamous non-small cell lung cancer
AUTHORS	Ozawa, Yuichi; Sugimoto, Takeya; Azuma, Yuichiro; Harutani, Yuhei; Yoshikawa, Takanori; Yamamoto, Nobuyuki; Kanai, Kuninobu

VERSION 1 – REVIEW

REVIEWER	Sung Yong Lee Korea University Guro Hospital, Republic of Korea
REVIEW RETURNED	20-Mar-2020

GENERAL COMMENTS	It is considered to be an important study in evaluating the usefulness of the combined treatment of immune checkpoint inhibitor and chemotherapy at the old age. Based on these results, it is thought that it can affect the treatment pattern of the old age patients.

REVIEWER	Rafael Santana-Davila University of Washington. United States
REVIEW RETURNED	06-Apr-2020

GENERAL COMMENTS	i dont understand why the authors want to publish the fact that they are going to conduct a clinical trial. I agree with them that whether or not triple therapy is beneficial in the elderly population is an important one, and the results of this study should be published when the results are available. but i dont believe that a manuscript describing the protocol is warranted. Furthermore. this trial is a very small trial which endpoint is overall response rate which is not the most important questions, what is the question is whether patients live longer and this trial will not answer that. lastly the statistics are based on a keynote study that for unknown reasons had a poor response rate in the control arm. a better statistics would be to obtain a median response rate of other clinical trials using carboplation and pemetrexed or better yet a clinical trial of this patient population such as PMID 23434351

REVIEWER	Tao Chen Liverpool School of Tropical Medicine
REVIEW RETURNED	20-Apr-2020

GENERAL COMMENTS	Yuichi Ozawa etal proposed a single arm study to assess efficacy and safety of pembrolizumab combined with carboplatin and pemetrexed among elderly patients with advanced, non-squamous non-small cell lung cancer. Overall, the article is structured well. However, my concerns are indicated as below. 1, the primary outcome is not consistently defined and varies as "the proportion of response" and "the overall response rate". This is vital element in a trial and needs to be clear without doubts. 2, how was the outcome assessed? In a blind way? As this is single arm study, clearly there is no measurement to control bias from the patient or investigator. However, it will be crucial to control the bias from outcome assessors. 3, the sample size is calculated by comparing the expected rate of triple regimes with the reference rate from CBDCA/PEM. However, this study is not to replicate the subgroup result from KEYNOTE189. The authors need to explain: a) Is 20% the minimal efficacy in the current practice? 2) what is the precision of the estimation? Like the lower boundary of the confidence interval. It seems the author only consider the difference of the point estimates. 4, It will be informative to have a Gantt chart to illustrate the whole picture of this study.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

We thank you very much for appreciation for our trial and protocol.
As reviewer1 pointed out, we believe our current prospective study have clinical significance in evaluation of the additional efficacy and safety of pembrolizumab when combined with CBDCA and PEM in elderly, chemo-naïve, patients with non-sq.

Reviewer: 2

We thank you for your suggestive opinions.
Your comments were helpful in reassessing the value and significance of this study and protocol.

#1: I agree with them that whether or not triple therapy is beneficial in the elderly population is an important one, and the results of this study should be published when the results are available. but i dont believe that a manuscript describing the protocol is warranted.

First of all, we would like to thank you for your understanding of the clinical significance

of the current trial focusing clinical significance of the triplet regimen in the elderly.

About the significance of papers that explain protocols, although clinical trials are tend to be evaluated based on only results, the hypotheses, rationale, and methodologies are also of great value. We believe that it will be important to share these and discuss for the further development clinical trials in future.

#2: this trial is a very small trial which endpoint is overall response rate which is not the most important questions, what is the question is whether patients live longer and this trial will not answer that.

As review 2 pointed out, setting response rate as primary endpoint is not as reliable as survival time to evaluate clinical benefit. However, it is clear in the era of ICI that survival time in patients with a response is longer than that in patients without a response, and it is sufficiently valuable as an alternative index for the efficacy. Furthermore, to set survival time as primary endpoint, hundreds of cases will be required, and it is difficult to carry out at least in recent situation. Therefore, in terms of the purpose of this study, which is to evaluate clinical benefit of the triplet regimen in elderly, it is considered appropriate to define response rate as primary end point.

#3: the statistics are based on a keynote study that for unknown reasons had a poor response rate in the control arm. a better statistics would be to obtain a median response rate of other clinical trials using carboplation and pemetrexed or better yet a clinical trial of this patient population such as PMID 23434351

As indicated, the response rate of CDDP / PEM was 30.6% in 862 cases (median age: 61 years, range: 28.8 - 83.2) in the phase 3 study in 2008 (GV Scagliotti, et al. JCO, 2008), and in the report you cited, which enrolled only elderly (age ≥70, n = 62), it was reported that response rate was 28.6%. However, those studies assessed response based on RECIST 1.0 although we use RECIST 1.1. Because the PR confirmation required by the RECIST 1.1, not RECIST 1.0, it would be appropriate to estimate the response rate of CBDCA/PEM to be less than previous report with RECIST 1.1.
Furthermore, in the PMID 23434351 study, the 28.6% was calculated in 44 "qualified" patients who were treated at least once and evaluated for efficacy at least once, and the response rate in ITT evaluation was just 22.6%. In addition, it also may be important that PMID 23434351 study included people over 70 years of age, whereas this study included only over 75 years of age.

In Keynote189, the response rate in platinum plus pemetrexed was 18.9%, which evaluate 206 patients in double-blind using RECIST1.1. In this study, the efficacy evaluation was set (at 6, 12 weeks, and then every 9 weeks) and planned to be evaluated using RECIST 1.1 with reference to KEYNOTE 189. Considering that CDDP, which have slightly high tumor shrinking activity, was not allowed and enrolling only elderly (≥75) in this study, it would be difficult to exceed response rate of 18.9%.

Based on these, we think that it is acceptable to make a hypothesis with a threshold of 20% (and an expected value of 45%) in this study.
We also added a description for size calculation and cited the suggested paper (R

Gervais, et al. Lung Cancer, 2013) on page 14 line (line 217 – 220).

Reviewer: 3

We thank you for your insightful comments, which we believe have helped us to improve our manuscript. Our specific responses to your points are as follows:

1, the primary outcome is not consistently defined and varies as "the proportion of response" and "the overall response rate". This is vital element in a trial and needs to be clear without doubts.

Thank you for pointing out the important points.
We have unified those to "the overall response rate".
The corrected part is in red.

2, how was the outcome assessed? In a blind way? As this is single arm study, clearly there is no measurement to control bias from the patient or investigator. However, it will be crucial to control the bias from outcome assessors.

We agree that it is important to evaluate response without prejudice.

In this study, response will be evaluated by each investigators. However, to avoid prejudice, EDC collects the size of target lesions, condition of non-target lesions, and presence/absence of new lesions in all CT images taken during the trials and evaluations will be confirmed by an expert committee.
We added a description on Page 13 (line 202 – 203)

3, the sample size is calculated by comparing the expected rate of triple regimes with the reference rate from CBDCA/PEM. However, this study is not to replicate the subgroup result from KEYNOTE189. The authors need to explain: a) Is 20% the minimal efficacy in the current practice? b) what is the precision of the estimation? Like the lower boundary of the confidence interval. It seems the author only consider the difference of the point estimates.

a) We deeply appreciate for your important pointing out. As indicated, the response rate of CDDP / PEM was 30.6% in 862 cases (median age: 61 years, range: 28.8 - 83.2) in the phase 3 study in 2008 (GV Scagliotti, et al. JCO, 2008). Other paper which enrolled only elderly (age ≥70, n = 62) showed response rate of 22.6% (ITT cases) to 28.6% (qualified cases). (R Gervais, et al. Lung Cancer, 2013)

One of the reasons why the response rate of Keynote 189 is lower than previous reports may be difference of criteria for response. RECIST 1.0 is used in the above 2 studies whereas it was RECIST 1.1 in Keynote 189, and the PR confirmation required by the RECIST 1.1, not RECIST 1.0.
In this study, the efficacy evaluation was set (at 6, 12 weeks, and then every 9 weeks) and planned to be evaluated using RECIST 1.1 with reference to Keynote 189.
Considering that CDDP, which have slightly higher tumor shrinking activity, was not

allowed in this study and enrolling only elderly (≥75, not ≥70) in this study, we think that it would be difficult to exceed response rate of 18.9%.

Based on these, we think that it is acceptable to make a hypothesis with a threshold of 20% (and an expected value of 45%) in this study.

To clarify that this is not to replicate the subgroup result from KEYNOTE189, we added a description for size calculation and cited the paper mentioned above (R Gervais, et al. Lung Cancer, 2013) on page 14 line (line 217 – 220).

b) For the calculation of case numbers in this study, we only consider the difference of the point estimates. Considering the differences between previous studies and current study as mentioned above, we believe that the current number of cases are sufficient enough to evaluate an efficacy of the triplet regimen.

4, It will be informative to have a Gantt chart to illustrate the whole picture of this study.

As suggested, we made a Gantt chart, which is shown as Figure 2 on page 12.
we also believe this has made it easier to understand the entire image of the study.