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)	Pembrolizumab Monotherapy for Non-Small Cell Lung Cancer (NSCLC): Can Patient Stratification be Improved in the UK Tayside Population? A Retrospective Cohort Study
AUTHORS	Mander, Emily; Merrick, Christopher; Nicholson, Hugh; Lord, Hannah; Ferguson, Michelle; Smith, Gillian

VERSION 1 – REVIEW

REVIEWER	De Castro, Jr, Gilberto Santa Maria delle Croci Hospital, Department of Oncology
REVIEW RETURNED	13-Aug-2023

GENERAL COMMENTS	In the study, the authors explore a retrospective cohort from a single institution in Scotland to investigate outcome modifiers for monodrug therapy with pembrolizumab in patients with non-small cell lung cancer. The decision to stratify patients according to the objective of radiotherapy treatment was certainly interesting and valuable. Although well planned and carried out with mostly solid statistics, the study suffers from being unicentric and having a small number of patients. Furthermore, some questions should be addressed by the authors. 1) Please consider shortening the Background section. 2) In the methods section the the authors mention that's patients who died after one cycle of pembrolizumab were excluded. The practice is concerning because it may mask patients with a fast disease progression. A sensitivity analysis should be run to better understand the effects of this decision. 3) An exploratory analysis with logistic regression could have been conducted to evaluate the effect of a continuous variable such as TPS as a predictive and prognostic tool, leveraging the disadvantage of being unicentric as a feature, by having a mostly uniform team and using a single antibody. 4) Would separating light smokers from heavy smokers in the analysis have made a difference? 5) Regarding the analysis of TPS and smoking status and survival: smoking status can be a confounder for TPS and vice-versa, as smoking can be positively correlated to TPS (e.g. PMID 30305940). An exploratory analysis looking at smoking status per TPD stratum should be very interesting as well. 6) As most patients who received radiotherapy have received it with palliative intent, not SABR, it is possible that this population had a higher burden of disease, making radiotherapy simply a surrogate indicator of higher burden instead of treatment. 7) Having most of the body of evidence doing the analysis with TPS 90% as a threshold should warrant the use of this measure at least in an exploratory analysis.

REVIEWER	Su, Chunxia Tongji University Affiliated Shanghai Pulmonary Hospital
REVIEW RETURNED	23-Aug-2023

GENERAL COMMENTS	This article primarily investigates the impact of PD-L1 TPS≥80%, smoking status, and prior or within-2-months-of-treatment radiotherapy on the PFS of NSCLC patients receiving pembrolizumab monotherapy. However, there are already existing articles that discuss the effect of PD-L1 TPS stratification on the efficacy of immune checkpoint inhibitors (ICIs) for NSCLC. The novelty of this article is limited, and the content is relatively straightforward. Here are some specific suggestions: （1）Were all the patients included in this study receiving first-line pembrolizumab monotherapy? If not, please provide a clear description. （2）Is there a difference in treatment efficacy among patients with metastases (such as brain metastases) based on different PD-L1 TPS? I am also interested in exploring this aspect. （3）Please specify the PD-L1 detection antibody used. Were all of them Dako 22C3 antibodies? Because different antibodies could yield different results. （4）In the discussion, it is mentioned that both US and Japanese studies utilized a higher PD-L1 TPS threshold (≥90%) for patient stratification. Why did the authors choose to set their PD-L1 stratification at 80%? What is the rationale behind this decision? （5）As a retrospective study, the sample size is limited to 100 cases. Considering that there are more approved ICIs for NSCLC, could the authors consider including more varieties of ICIs to expand the sample size?

REVIEWER	Frost, Nikolaj Charite Universitatsmedizin Berlin
REVIEW RETURNED	28-Aug-2023

GENERAL COMMENTS	Mander and colleagues provide a single-center retrospective cohort analysis of 100 patients with PD-L1 high NSCLC treated with pembrolizumab as palliative first-line treatment. Given rigid in- and exclusion criteria in registrational phase III trials the generation of real-world evidence for questions not addressed in the respective trials is key. This investigation deals with important questions, the clinician is faced with in everyday clinical practice. These include a potential correlation of PD-L1 expression level and treatment efficacy and the impact of smoking history and radiotherapy. The endpoint chosen was PFS, which is appropriate for the questions raised. However, I wonder, why the authors only used progression as a definition and not additionally death from any cause (whichever might occur first) as usually. Second, real-world PFS (this means imaging without standardized RECIST assessment) might not correlate with PFS from clinical trials. The authors should comment this. First, outcome according to PD-L1 expression levels was investigated (50-79% vs. 80-100%) and did not show any difference. It is not clear for me why this cut-off was chosen, as retrospective (e.g. Aguilar et al., Ann Oncol 2019, ref. 39) and prospective data (e.g. Sezer et al., Lancet 2021) demonstrated a correlation of every percentile of PD-L1 expression and a separation of survival curves in the PD-L1 highest group.

	Second, survival differences according to receipt of radiotherapy and its timing were assessed. RT not only was not better than pembrolizumab alone but in contrast was associated with worse survival. In a non-randomized setting, prone to a large number of bias, there is a high probability that RT is only a surrogate for more advanced and aggressive disease, as the authors state in the discussion. I do not agree with the estimation that the Pembro-RT trial did not show any differences. Of course, statistical significance was missed (due to the limited number of patients included), but there was a clear trend in favor of additional RT. Table 1 in the supplement mentions “mutations” with regard to ALK and ROS1, this should be replaced with “re-arrangements” or “fusions”. The authors should also state, whether there were any significant differences with regard to baseline characteristics. Finally I suggest to add a table with the respective forest plots for the Cox regression analyses. To conlude, this is a welcome manuscript adding important information but I suggest to address the mentioned points, to moderate the key message a little bit (“stratification of PD-L1 TPS is not warranted) and to add a limitations paragraph at the end of the manuscript.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 comments:

1. Please consider shortening the Background section.
We have shortened our Background section as requested.
2. In the methods section the authors mention that patients who died after one cycle of pembrolizumab were excluded. The practice is concerning because it may mask patients with a fast disease progression. A sensitivity analysis should be run to better understand the effects of this decision.
There were an additional 7 patients in our original patient cohort who died after one cycle of pembrolizumab treatment – as PFS was our primary analysis end point, it was not possible to include these patients in extended analysis as disease progression had not been confirmed by CT scan. To address this point, we have re-run our OS survival analysis, and confirm that the inclusion of these patients did not change our conclusions. PD-L1 expression in the extended dataset again did not influence OS (HR=0.108, p=0.742), and our smoking and radiotherapy conclusions were similarly not affected. We have now commented on this additional analysis in our revised Discussion (page 11 – please note that we refer throughout our response to page numbers in the tracked changed copy of our revised manuscript).
3. An exploratory analysis with logistic regression could have been conducted to evaluate the effect of a continuous variable such as TPS as a predictive and prognostic tool, leveraging the disadvantage of being unicentric as a feature, by having a mostly uniform team and using a single antibody.
We agree that logistic regression analysis to investigate the potential to use TPS as a predictive or prognostic tool would be important, if we had access to more quantitative TPS data. Our current “real world” clinical dataset is unfortunately not sufficiently linear or quantitative to support the suggested logistic regression analysis, with assessment by more than one pathologist leading to further variation in TPS classification. We had already highlighted this point in our Discussion, and have added further text (Discussion, page 11) to highlight that this is an important priority for future studies.
4. Would separating light smokers from heavy smokers in the analysis have made a difference?

Again, we agree that this is a very interesting area for future study – but we unfortunately don't have access to quantitative smoking data, beyond the current/previous smoker data we have reported. We have already discussed the limitations of self-reported smoking data, and have now more clearly highlighted that more quantitative assessment of smoking data, for example following measurement of serum cotinine levels, is a priority for future studies (Discussion, page 12).

5. Regarding the analysis of TPS and smoking status and survival: smoking status can be a confounder for TPS and vice-versa, as smoking can be positively correlated to TPS (e.g. PMID 30305940). An exploratory analysis looking at smoking status per TPS stratum should be very interesting as well.

We agree with the reviewer, but the suggested analysis would require access to more quantitative TPS and smoking data – we have extended our Discussion as described above to highlight this point.

6. As most patients who received radiotherapy have received it with palliative intent, not SABR, it is possible that this population had a higher burden of disease, making radiotherapy simply a surrogate indicator of higher burden instead of treatment.

We agree that individual patients will have had different burdens of disease, and that this is likely to have influenced treatment decisions. Our Cox regression analysis was designed to address potential confounding variables, where we confirmed that the radiotherapy associations we report were independent of tumour stage and histology.

7. Having most of the body of evidence doing the analysis with TPS 90% as a threshold should warrant the use of this measure at least in an exploratory analysis.

36 patients had TPS□90% - we were initially concerned that we would be significantly under-powered using a TPS□90% threshold, but have now completed this exploratory analysis. We did not find any significant influence of TPS on either PFS (HR=0.061, p=0.805) or OS (HR=0.001, p=0.976), when a TPS□90% threshold was used. We have added this information to Results, Section 2.

Reviewer 2 comments:

1. Were all the patients included in this study receiving first-line pembrolizumab monotherapy? If not, please provide a clear description.

Not all patients received first-line pembrolizumab monotherapy. We apologise for this omission and have now included more detailed patient demographics in our revised Table 1.

2. Is there a difference in treatment efficacy among patients with metastases (such as brain metastases) based on different PD-L1 TPS? I am also interested in exploring this aspect.

The majority of our patients were Stage 4, assessed according to standard TNM criteria, with only 30 patients having no evidence of metastatic disease at diagnosis. Many of our Stage 4 patients were however reported as node positive, but with no specific metastatic site recorded, so it is unfortunately not possible to meaningfully further sub-divide our dataset in this patient cohort. In exploratory analysis, however, we have further sub-divided our patient cohort according to the presence or absence of metastatic disease. Not surprisingly, the presence of metastatic disease negatively influenced overall survival (HR=14.741, p<0.001) – progression free survival was not formally significant (HR=3.243, p=0.072), although a clear trend was observed. PD-L1 TPS did not influence PFS (metastatic HR=0.222, p=0.638, non-metastatic HR=1.947, p=0.163) or OS (metastatic HR=0.880, p=0.348, non-metastatic HR=1.676, p=0.195) in either the metastatic or non-metastatic group.

3. Please specify the PD-L1 detection antibody used. Were all of them Dako 22C3 antibodies? Because different antibodies could yield different results.

As described in our Methods section, the Dako 22C3 antibody was used for PD-L1 detection in all patients.

4. In the discussion, it is mentioned that both US and Japanese studies utilized a higher PD-L1 TPS threshold (≥90%) for patient stratification. Why did the authors choose to set their PD-L1 stratification at 80%? What is the rationale behind this decision?
Please see our comments above in response to Reviewer 1, point 7.
5. As a retrospective study, the sample size is limited to 100 cases. Considering that there are more approved ICIs for NSCLC, could the authors consider including more varieties of ICIs to expand the sample size?
Although additional ICIs have subsequently been licensed, pembrolizumab was the only ICI approved in Scotland for first line monotherapy in advanced NSCLC in patients with PD-L1 TPS ≥50% with no EGFR mutations or ALK translocations, during the period of data collection for our study.

Reviewer 3 comments:

1. The endpoint chosen was PFS, which is appropriate for the questions raised. However, I wonder, why the authors only used progression as a definition and not additionally death from any cause (whichever might occur first) as usually.
We have reported PFS as our primary outcome variable, but have already included OS (death from any cause) in our Supplementary information.
2. Real-world PFS (this means imaging without standardized RECIST assessment) might not correlate with PFS from clinical trials. The authors should comment this.
This is an important point, which was highlighted by Reviewer 2 as a particular strength of our study. We have already commented on this point in our Discussion (page 13).
3. First, outcome according to PD-L1 expression levels was investigated (50-79% vs. 80-100%) and did not show any difference. It is not clear for me why this cut-off was chosen, as retrospective (e.g. Aguilar et al., Ann Oncol 2019, ref. 39) and prospective data (e.g. Sezer et al., Lancet 2021) demonstrated a correlation of every percentile of PD-L1 expression and a separation of survival curves in the PD-L1 highest group.
Please see our comments above – our decision to initially stratify patients based on TPS□80% was partly based on statistical power, as only one third of patients had TPS□90%, and also for consistency with previous literature which classifies TPS□50% as high and TPS□70% as very high.
4. Second, survival differences according to receipt of radiotherapy and its timing were assessed. RT not only was not better than pembrolizumab alone but in contrast was associated with worse survival. In a non-randomized setting, prone to a large number of bias, there is a high probability that RT is only a surrogate for more advanced and aggressive disease, as the authors state in the discussion. I do not agree with the estimation that the Pembro-RT trial did not show any differences. Of course, statistical significance was missed (due to the limited number of patients included), but there was a clear trend in favor of additional RT.
We chose to describe previous significant studies only when formal statistical significance was achieved, but have added additional text to our Discussion section (page 12) to highlight the trend suggested by the Reviewer.
5. Table 1 in the supplement mentions “mutations” with regard to ALK and ROS1, this should be replaced with “re-arrangements” or “fusions”.
We apologise for this omission and have corrected Table 1 and similar references within our manuscript text.
6. The authors should also state, whether there were any significant differences with regard to baseline characteristics.
We apologise for this omission and have added pairwise comparisons to the demographic data presented in Table 1, with associated methods information now included in Methods, Section 7 Statistical Analysis. There were no significant differences in baseline characteristics, comparing patients with PD-L1 TPS 50-79% and □80%.
7. Finally I suggest to add a table with the respective forest plots for the Cox regression analyses.

Forest plots are more commonly used to illustrate the output of meta analysis studies – statistical reporting of our Cox regression analysis summarises hazard ratios, with associated standard errors and p-values, and does not include the 95% confidence interval estimates required for Forest plot illustration. We have however summarised our Cox regression analysis in more detail in Supplemental Material Appendix 2.

We hope that our response satisfactorily addresses each of the points raised by the Editor and Reviewers and attach a tracked change and final revised copy of our manuscript.

VERSION 2 – REVIEW

REVIEWER	Frost, Nikolaj Charite Universitatsmedizin Berlin
REVIEW RETURNED	21-Sep-2023

GENERAL COMMENTS	The revised manuscript adressed nearly all questions. However, definition of PFS seems not correct to me (point 1). In general, PFS is defined as time to radiologically disease progression or death from any cause, whichever occurs first. In the defintion from Mander et al., patients having dies would have been censored without radiographically defined progression. The authours should explain why this distinct definition was chosen or should change it (which I would recommend).

VERSION 2 – AUTHOR RESPONSE

We agree with Dr Frost's definition of PFS which applies to all patients in our cohort, with the exception of the 7 patients we chose to exclude from our survival analysis. As we explained in our previous response to the reviewers, these 7 patients were diagnosed with advanced stage disease, and died following only a single pembrolizumab treatment, before it was possible to assess radiological response. We therefore chose to exclude these patients, as it was not possible to accurately estimate PFS in the absence of radiological assessment. We have already described these excluded patients in Figure 1, but have now added additional information to our Methods Section 2 to more clearly explain our exclusion strategy. Although we provided additional analysis in our previous response to the reviewers to confirm that our overall survival conclusions were not influenced by the exclusion of these patients, our preference would be to continue to exclude them from our analysis, to allow the correct definitions of PFS and OS to be maintained for all included patients.