Peer Review File

Article Information: https://dx.doi.org/10.21037/tlcr-21-815

Reviewer A

Comment 1.1 (C1.1) This is a single center retrospective study evaluating baseline tumor size as a predictor of outcomes from ICI. The authors evaluated max BTS (maximum target lesions' longest diameter) and total BTS (sum of target lesions' longest diameters) and report that max BTS was a/w worse PFS in pts treated with ICI-monotherapy but not in ICI+chemo, and that patients with larger tumors (BTS > 5cm) had worse outcomes treated with ICI mono compared to ICI chemo.
The finding that tumor burden correlates with outcomes is of course not novel. The major concern I have in this study is generalizability. The study was 80% men; over 35% of patients in ICI mono had PS 2 (or higher), compared to <3% in ICI + chemo arm…and there were 20% of patients in ICI+chemo with unknown PD-L1.

Response 1.1 (R1.1)
We appreciate these insightful comments. As described in the **Results** section (**page6, para1**), we had excluded the patients with driver mutations such as EGFR mutation, which was a common driver mutation among never-smoker Asian women. It was considered to lead to the result that 80% of this cohort was men. As suggested, we have added the important information in the limitation section of the Discussion section (**page 15-16, para10**) as follows;

Change in the text:
Most of our cohort was men because we excluded the patients with driver mutations such as EGFR mutation, which was a common driver mutation among never-smoker Asian women.

As shown in the limitation of the **Discussion** section (**page 15-16, para10**), we agree that the patients with poor PS tended to receive ICI-mono. If the PD-L1 expression is unknown, ICI-chemo tended to be the prefeed choice because it had been approved regardless of PD-L1 status. Our multivariate analysis, including sex, PS, and PD-L1 expression, could help minimize the potential for these confounding biases. We have clarified this point in response 1.2.

C 1.2 The authors state that "in patients with max BTS ≥ 50 mm, the multivariable analysis showed that the choice of ICI chemo vs. ICI-mono was an independent predictive factor for PFS." The inference here is that patients with high tumor volume should receive more aggressive care – which is often done clinically. The major problem as I see it is that PD-L1 expression was associated with PFS and OS in every single analysis. Therefore the true driver of outcomes is likely PD-L1 expression and not max BTS. The ICI monotherapy arm also had generally worse PS and was older, thus skewing the results and limiting generalizability.

R1.2 We appreciate the opportunity to clarify this point. As shown in the limitation of the **Discussion** section (**page 15-16, para10**), we agree with the reviewer that the patients with

aggressive disease tended to receive ICI-chemo. In addition, we admit the strong association of PD-L1 expression with PFS, OS in our analysis might confound our result. However, previous studies showed that PD-L1 expression also had limitations as a predictive biomarker in NSCLC. For example, the PD-L1 expression seemed to be lower in primary tumor samples than in metastases and the extent of PD-L1 expression in metastases varies by anatomic site, and there was lower PD-L1 expression in resection specimens compared to biopsy samples (1). As suggested, we have discussed this issue in the limitation section of the **Discussion** section (**page 15-16, para10**).

Change in the text:

Before: Although the baseline clinical characteristics of the patients in the ICI-mono and ICI-chemo groups were moderately well-balanced, we acknowledged the existence of a selection bias that patients with aggressive status, younger age, better PS, and no significant medical comorbidities were more likely to receive ICI-chemo.

After: Although the baseline clinical characteristics of the patients in the ICI-mono and ICI-chemo groups were moderately well-balanced, we acknowledged the existence of a selection bias that patients with aggressive status, younger age, better PS, and no significant medical comorbidities were more likely to receive ICI-chemo and patients with indolent status, older age, poorer PS, and many medical comorbidities were more likely to receive ICI-mono. The strong association of PD-L1 expression with clinical outcomes might confound our result; however, previous studies have shown that PD-L1 expression has limitations as a predictive biomarker in NSCLC (1). Our evaluation of the max BTS with multivariate analysis, including number of organs involved, age, PS, and PD-L1 expression, could help minimize the potential for confounding bias.

C1.3 When it comes down to analysis, for the PDL1 high subset we were left with a comparison of 24 vs 44 patients. Given how different the outcomes were from other data sets (even real world data ie PMID: 31435660 where OS wasn't reached and here mOS was only 15 months), I worry that these small numbers preclude meaningful interpretation. For the ICI + chemo subset, the PD-L1 high comparison was only 13 vs 5 pts (ie Fig 3 D).

R1.3 We appreciate these helpful suggestions. Our hospital having many older patients with comorbidities or low income, the median OS in our patients with PD-L1 TPS≥ 50% was reasonable compared to IMpower 110 (PD-L1 ≥ 50%; median OS, 20.2 months). We consider it difficult to compare our hospital's OS with the special cancer centers, including Dana-Farber Cancer Institute and Massachusetts General Hospital Cancer Center, as you cited (2). The real-world data in the Netherlands showed that median OS in patients with PD-L1 ≥ 50% who received ICI-mono was around 15 months, which was similar to our result (3). We agree the comparison of 13 vs. 5 patients with PD-L1 ≥ 50% in the ICI-chemo group might demonstrate the importance of PD-L1; however, statistical power is inadequate due to the small number. As mentioned above in the response 1.2, we have added the information that PD-L1 could be a confounding factor.

C1.4 The comparison that should be done is focusing only on PDL1 > 50% subset vs PDL1 < 50%. Patients with PDL1< 50% that are offered ICI monotherapy is usually because they are either not a candidate for chemotherapy in the eyes of their treating oncologist, or patient preference. As the authors state, the SOC would be combination therapy (ICI chemo or other ICI combinations) for these patients.

R1.4 We appreciate the reviser's point**.** Although we agree with the importance of comparations stratifying patients with PD-L1 ≥ 50%, the statistical power of the current study is inadequate due to the small number of our subgroup. Also, as we mentioned in the response 1.1 and 1.2, PD-L1 was not a complete biomarker. Our multivariable analysis could reduce the bias from this confounding.

C1.5 Also, no rationale is given for the inclusion of stage III patients – these should be excluded entirely. There is not enough justification for the "recurrence" arm which was over 50 patients – were these locally recurrent or widely metastatic? How soon after treatment did these patients recur? These patients made up 30% of the ICI mono cohort (vs 10% in ICI chemo), which may also be skewing results.

R1.5 We thank the reviewer for the careful review of the manuscript. We excluded the recurrence of stage III patients after chemoradiotherapy because a considerable number of their patients received durvalumab according to the PACIFIC regimen. The inclusion of patients with the previous ICI treatment was considered inappropriate for analyzing patients with first-line ICIs. We have revised this part to convey a clear message in the **Results** section (**page6, para1**) as follows;

Change in the text:

Before: "Of the 191 potentially eligible patients, we excluded 23 with recurrence after chemoradiotherapy, one with EGFR/ALK/ROS1/BRAF alterations, and five clinical trial patients who received experimental regimens unapproved by FDA (Figure S1)."

After: "Of the 191 potentially eligible patients, we excluded 23 patients with recurrence after chemoradiotherapy because their patients included previous durvalumab treatment. Additionally, we excluded one patient with *EGFR*/*ALK*/*ROS1*/*BRAF* alterations and five clinical trial patients who received experimental regimens unapproved by FDA, thus no mutations in *EGFR*/*ALK*/*ROS1*/*BRAF* were identified in our cohort (Figure S1)."

We completely agree with the reviewer that there are various patterns of recurrence. We usually treat oligo-recurrence with local therapy. Thus, most 'recurrence' represented patients with widespread metastatic cancer or oligo-recurrence resistant to the local treatment. The former pattern was considered similar to stage IV. However, the detail of their recurrent patterns was different for each patient. Additionally, there was no clear consensus on how to classify these recurrent patterns. We consider it inappropriate to divide their patients into their subgroups. As suggested, we have added the important information in the limitation section of the Discussion section (**page 15-16, para10**) as follows;

Change in the text:

Although the recurrent at staging represented patients with widespread metastatic cancer or oligo-recurrence resistant to local treatment, the detail of their recurrent patterns was different for each patient.

C1.6 There is no mention of mutational status for patients which should be provided and evaluated as these may inform second line choice of treatment and therefore better outcomes.

R.1.6 We apologize for the confusion. Indeed, we excluded patients who had *EGFR*/*ALK*/*ROS1*/*BRAF* alterations, as we had described in the **Results** section (**page6, para1**) and **Figure S1**. We have revised this part to carry a clear message as mentioned above in the response 1.5.

Minor points:

C2. In patients with a PD-L1 TPS ≥ 50%, the median PFS was 4.5 months (95% CI, 2.1–10.8) in the 234 ICI-mono group and 15.7 months (95% CI, 7.7–NA) in the ICI-chemo group. In patients with a PD-L1 TPS ≥ 50%, the median OS was 15.6 months (95% CI, 9.3–34.7) in the ICI-mono group

This raises concern for me given that this patient population generally does well, therefore the worse outcomes seen in this study should be explored and a rationale offered. This is a concern for generalizability.

R2. I appreciate the reviewer's point, which we consider is also essential. In max BTS ≥ 50 mm, there were a few patients with PD-L1 TPS ≥ 50% who received ICI-chemo (n=9), which was difficult to show a conclusive result statistically. It might have a selection bias that better PS, younger age, and no significant medical comorbidities were more likely to receive ICI-chemo. We have changed the limitation in the **Discussion** section (**page 15-16, para10**).

Change in the text:

Before: First, the sample size and was relatively small, especially in PD-L1 TPS ≥ 50%.

After: First, the sample size was relatively small, especially in PD-L1 TPS ≥ 50% with lower statistical power.

C3. For the max BTS lesion, perhaps I missed this, but were these all pulmonary lesions? Or could it just be the largest lesion overall? The location should be provided.

R3. We thank the reviewer for the careful review of the manuscript. The max BTS lesion included locations other than pulmonary lesions. As requested, we have added the table to the supplement as follows;

New Table S1. Site of the largest target lesions

Site of the largest target lesions, n (%)	ICI monotherapy (n=80)	ICI + Chemo (n =79)	P
Primary lesion	57 (70.4)	55 (69.6)	0.135
Bone	4 (5.0)	7 (8.9)	
Intrathoracic lymph node	3 (3.7)	5 (6.3)	
Pleura	6 (7.4)	2 (2.5)	
Brain	0 (0.0)	3 (3.8)	
Adrenal gland	4 (5.0)	2 (2.5)	
Lung (metastasis)	0 (0.0)	2 (2.5)	
Others	4 (5.0)	2 (2.5)	
Liver	0 (0.0)	1 (1.3)	
Extrathoracic lymph node	3 (3.7)	0 (0.0)	

Also, we have added this information to the **Result section** as follows (**page 7, para1**);

Change in the text: The most common site for the largest target lesions was a primary pulmonary lesion.

C4. Any cases of pseudo progression?

R4. We thank the reviewer for the insightful comment. There might be some pseudoprogression cases. However, there is currently no consensus to define pseudoprogression and they are rare events that have been reported in 2% to 8% of patients with advanced NSCLC treated with ICIs (4). We therefore think these cases would not change our principal analysis.

C5. The addition of data from another site would strengthen this considerably and mitigate some of the potential confounding variables.

R5. We fully agree with the review in that the data from another site would improve generalizability. Unfortunately, we do not have these data. In the limitations of **Discussion** section (**page 15-16, para10**), we have acknowledged this issue.

Reference

1. Brueckl WM, Ficker JH, Zeitler G. Clinically relevant prognostic and predictive markers for immune-checkpoint-inhibitor (ICI) therapy in non-small cell lung cancer (NSCLC). BMC Cancer. 2020 Dec 3;20(1):1185.

2. Aguilar EJ, Ricciuti B, Gainor JF, Kehl KL, Kravets S, Dahlberg S, et al. Outcomes to first-line pembrolizumab in patients with non-small-cell lung cancer and very high PD-L1

expression. Annals of Oncology. 2019 Oct 1;30(10):1653–9.

3. Cramer-van der Welle CM, Verschueren MV, Tonn M, Peters BJM, Schramel FMNH, Klungel OH, et al. Real-world outcomes versus clinical trial results of immunotherapy in stage IV non-small cell lung cancer (NSCLC) in the Netherlands. Sci Rep. 2021 Mar 18;11(1):6306.

4. Ferrara R, Caramella C, Besse B, Champiat S. Pseudoprogression in Non–Small Cell Lung Cancer upon Immunotherapy: Few Drops in the Ocean? Journal of Thoracic Oncology. 2019 Mar 1;14(3):328–31.

Reviewer B

Although it is understood that increase tumor size and the increasing number organs may lead to and cancer outcomes, but the objective data on this postulation is lacking. Authors have done a great job to objectify this hypothesis and have done a great and comprehensive analysis.

C: We thank the reviewer for the positive comment.

Reviewer C

I think this is an interesting look at a common clinical issue. Although the trial is relatively small, the authors acknowledged their limitations and included future directions.

C: We thank the reviewer for the excellent comment.

Reviewer D

The authors presente an interesting retrospective work about the efficacy of ICI+/- CT according to the tumor size. The current data endorse previous hyopothesis that ICI monotherapy is more suitable for small and indolent tunors, whereas CT+ICI effiacy occurss regardless of the tumor burden. It may help to decide in our daily clincial practice for decidign what patients with high PDL1 expression are more suitable for monotehrapy vs combos. It is a well written manuscript and I only have one minor comment:
In the discussion authors could include the data from KN189 reprted by Gadgeel at AACR 2021 that efficacy of chemo+ICI occurred regarldes more or less than 3 metastatic sites, supporting their observation (DOI:10.1158/1538-7445.AM2021-442)

C: We thank the reviewer for the positive comment and helpful paper. As suggested, we have added this information to the **Discussion section** (**page 15, para7**):

Change in the text:

Before: "Our findings suggest that BTS might be evaluated in future ICI trials and integrated into clinical practice to provide treatment expectations when choosing between ICI-mono vs. ICI-chemo**."**

After: "A post-hoc efficacy analysis of KEYNOTE-189 showed ICI-chemo showed consistent clinical benefit regardless of BTS(5). These findings suggest that BTS might be evaluated in future ICI trials and integrated into clinical practice to provide treatment expectations when choosing between ICI-mono vs. ICI-chemo (30)**."**

Reference

30. Gadgeel S, Grey J, Rizzo MT, Peterson P, Kim J, Rodríguez-Abreu D. Abstract 442: Pemetrexed and platinum plus pembrolizumab in patients with metastatic non-squamous non-small cell lung cancer by tumor burden at baseline: A post-hoc efficacy analysis of KEYNOTE-189. Cancer Res. 2021 Jul 1;81(13 Supplement):442–442.