nature portfolio

Peer Review File

Nivolumab plus ipilimumab for chemotherapy-refractory metastatic castration-resistant prostate cancer: results from the randomized portion of the phase 2 CheckMate 650 trial

Corresponding Author: Dr Padmanee Sharma

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The author has addressed most of my questions. The only question I have is “Why are the false discovery rate threshold chosen to be 0.0001?”
In the author’s response, “P-values were adjusted by the Bonferroni correction to maintain a stringent family-wise error rate. We set the adjusted p-value threshold at 1×10^{-4} minimizing false positives across >10,000 tests—and further required an average fold change > 2.0 to ensure that selected genes exhibit biologically meaningful differences.”
If the author used the Bonferroni correction stated: Then the threshold should be 0.05/10000=0.000005, which is much smaller than 1×10^{-4}. Please re-check the statistical analysis and make sure the inflate of type I error is controlled correctly.

Reviewer #5

(Remarks to the Author)
This review was solicited to assess the authors' response to the 3 prior reviews from the initial manuscript. As such, it is not a complete review of the paper.

This paper is a report of the results of a randomized phase 2 study assessing outcomes in metastatic prostate cancer of 2 regimens containing combinations of ipilimumab and nivolumab, ipilimumab alone or the control arm treated with cabazitaxel, a standard of care prostate cancer treatment. Secondarily, they found enrichment of certain cell types around blood vessels in what they defined as exceptional responders. Then they created a transcriptional profile based on those cell types and applied it to the TCGA database and found that the profile was associated with better outcomes.

For the most part, this is a straightforward study and was done very well. The authors addressed all the comments pertaining to the actual clinical part of the study.

Like the other reviewers, I had issues with the prognostic marker development. I agree with reviewer #4 that I find it unlikely that this prognostic marker will be valuable. More significant to me is the fact that the authors generated a cell biological-based prognostic marker based on response to immunotherapy and then used that to generate a relatively non-specific transcriptional profile that was used as a prognostic marker for outcomes in a dataset totally unrelated to immune therapies. While there is no problem with their findings, there is not much of a biological rationale for why their transciptional profile works.

That said, I think they have adequately addressed enough of the comments. I would like to see an explanation of their rationale for the transcriptional profile as a biomarker but I do not think it is absolutely necessary for publication.

Reviewer #6

(Remarks to the Author)
The authors present results from the CheckMate 650 trial in which patients with mCRPC were randomized to receive one of 4 treatment regimens with a goal of exploring alternative dosing of ipi/nivo checkpoint inhibitor therapy as well as benchmarking against cabazitaxel. While the efficacy results are disappointing, the authors have provided a fair and balanced discussion of the results and there is much that can be learned from the corrleative analyses.

A few comments:

1) The manuscript's impact is limited by lack of statistical design to compare between arms, or even to determine what would represent a "positive" outcome (meaning what amount of response defined by which parameter would be good enough to make a regimen considered active). the authors do mention that the sample size was chosen based on an expected ORR of approximately 30% in the N3I1 and/or N1I3 cohorts and that this would exceed what was expected from the CABA cohort. Given that the observed ORR was 9.3% for N3I1 and 19.5% in N1I3 cohorts, this would be considered a negative study.

2) The 5-gene perivascular immune niche correlative signature is interesting, particularly since the authors assessed it in TCGA and found better prognosis. this would seem to be a prognostic marker given that the majority of prostate cancer patients in TCGA most likely did not receive immunotherapy, even though the authors developed it in the context of immunotherapy-treated patients; there would need to be comparative assessment of its performance in CABA cohort to make a case for predictive ability and/or further validation. In fact, the inclusion of a 'control' arm of cabazitaxel really calls for study of correlatives in those patients as well, since clinical benchmarking could conceivably have done using relatively contemporary studies such as TheraP or CARD.

3) it is unclear why the authors suggest that liquid biopsy be used to guide patient selection in future efforts given the poor performance of liquid-based TMB assessment. Perhaps the authors intend that further liquid biopsy tests should be developed in concert with future immunotherapy studies, as they mention novel assays including extracellular vesicles.

4) would revise the following statement, since the data generated in this trial do not suggest that any of the immunotherapy regimens studied here are good enough to be applied in mCRPC - if there is to be a role for immune checkpoint inhibitor therapy in mCRPC beyond pembrolizumab's tumor agnostic indications, it will require better selective biomarkers. other immunotherapy agents, such as bispecific T cell engagers, are showing high response rates in unselected populations of mCRPC patients.

"Overall, these data add to the emerging body of evidence suggesting a potential role for immunotherapy-based regimens in the treatment of select patients with mCRPC"

Version 1:

Reviewer comments:

Reviewer #6

(Remarks to the Author)
The authors have fairly well addressed the remaining reviewer concerns.

REVIEWER COMMENTS

Original Reviewer #1 (biostats):

The author has addressed most of my questions. The only question I have is “Why are the false discovery rate threshold chosen to be 0.0001?”
In the author’s response, “P-values were adjusted by the Bonferroni correction to maintain a stringent family-wise error rate. We set the adjusted p-value threshold at 1×10^{-4} minimizing false positives across >10,000 tests—and further required an average fold change > 2.0 to ensure that selected genes exhibit biologically meaningful differences.”
If the author used the Bonferroni correction stated: Then the threshold should be 0.05/10000=0.000005, which is much smaller than 1×10^{-4}. Please re-check the statistical analysis and make sure the inflate of type I error is controlled correctly.

We appreciate the reviewer’s careful observation regarding our choice of statistical threshold. To clarify, we applied a stringent adjusted p-value cutoff of 1×10^{-4} rather than the conventional *P* value threshold of 0.05, using Bonferroni correction (equivalent to a raw *P* value threshold of 1×10^{-8} for 10,000 tests) to control type I error. In addition, we required an average fold change greater than 2.0 and selected the top five gene candidates. This approach was intentionally designed to minimize false positives while ensuring that the selected genes represent robust and biologically meaningful differences.

New Reviewer #5:

This review was solicited to assess the authors' response to the 3 prior reviews from the initial manuscript. As such, it is not a complete review of the paper.

This paper is a report of the results of a randomized phase 2 study assessing outcomes in metastatic prostate cancer of 2 regimens containing combinations of ipilimumab and nivolumab, ipilimumab alone or the control arm treated with cabazitaxel, a standard of care prostate cancer treatment. Secondarily, they found enrichment of certain cell types around blood vessels in what they defined as exceptional responders. Then they created a transcriptional profile based on those cell types and applied it to the TCGA database and found that the profile was associated with better outcomes.

For the most part, this is a straightforward study and was done very well. The authors addressed all the comments pertaining to the actual clinical part of the study.

Like the other reviewers, I had issues with the prognostic marker development. I agree with reviewer #4 that I find it unlikely that this prognostic marker will be valuable. More significant to me is the fact that the authors generated a cell biological-based prognostic marker based on response to immunotherapy and then used that to generate a relatively non-specific transcriptional profile that was used as a prognostic marker for outcomes in a dataset totally unrelated to immune therapies. While there is no problem with their findings, there is not much of a biological rationale for why their transciptional profile works.

That said, I think they have adequately addressed enough of the comments. I would like to see an explanation of their rationale for the transcriptional profile as a biomarker but I do not think it is absolutely necessary for publication.

We appreciate the feedback from the reviewer. We have clarified in the text (page 16): "… we derived a transcriptional signature from the top five differentially expressed genes across cell types identified by the CODEX assay using the single-cell prostate cancer reference[4], including more specifically defined cell types such as monocytes, endothelial cells, dendritic cells, CD4 naïve/central memory T cells, cytotoxic CD8 T cells, CD4 memory T cells, and CD8 memory T cells. This simplified signature was intended to capture the most abundant cell types identified by the spatial analyses."

New Reviewer #6:

The authors present results from the CheckMate 650 trial in which patients with mCRPC were randomized to receive one of 4 treatment regimens with a goal of exploring alternative dosing of ipi/nivo checkpoint inhibitor therapy as well as benchmarking against cabazitaxel. While the efficacy results are disappointing, the authors have provided a fair and balanced discussion of the results and there is much that can be learned from the corrleative analyses.

A few comments:

1) The manuscript's impact is limited by lack of statistical design to compare between arms, or even to determine what would represent a "positive" outcome (meaning what amount of response defined by which parameter would be good enough to make a regimen considered active). the authors do mention that the sample size was chosen based on an expected ORR of approximately 30% in the N3I1 and/or N1I3 cohorts and that this would exceed what was expected from the CABA cohort. Given that the observed ORR was 9.3% for N3I1 and 19.5% in N1I3 cohorts, this would be considered a negative study.

We thank the reviewer for this feedback. We have added this caveat to the Discussion (page 18):

"We note that although the study was not statistically powered to compare between cohorts, the observed ORR was below the expected ORR of approximately 30% in the N3I1 and/or N1I3 cohorts (based on a two-sided 95% CI of 16.8% to 45.2%, with the lower bound above the approximately 15% expected response rate for cabazitaxel based on the TROPIC study[8])."

2) The 5-gene perivascular immune niche correlative signature is interesting, particularly since the authors assessed it in TCGA and found better prognosis. this would seem to be a prognostic marker given that the majority of prostate cancer patients in TCGA most likely did not receive immunotherapy, even though the authors developed it in the context of immunotherapy-treated patients; there would need to be comparative assessment of its performance in CABA cohort to make a case for predictive ability and/or further validation. In fact, the inclusion of a 'control' arm of cabazitaxel really calls for study of correlatives in those patients as well, since clinical benchmarking could conceivably have done using relatively contemporary studies such as TheraP or CARD.

We agree with the reviewer that a comparison of the performance of this candidate biomarker in non-immunotherapy-treated cases would support its specificity. The small numbers of samples on this trial did not permit this comparative analysis. We have added the following to the Discussion (page 22):

“We acknowledge the small sample size for the biomarker analysis; these results are intended to be hypothesis-generating for subsequent validation in larger, prospective trials, including evaluation in patients treated with and without immunotherapy.”

3) it is unclear why the authors suggest that liquid biopsy be used to guide patient selection in future efforts given the poor performance of liquid-based TMB assessment. Perhaps the authors intend that further liquid biopsy tests should be developed in concert with future immunotherapy studies, as they mention novel assays including extracellular vesicles.

We thank the reviewer and have clarified the statement accordingly (page 23):

“Future clinical trials should incorporate these pretreatment tissue biomarkers alongside investigation of non-invasive liquid biopsy strategies to optimize patient selection. Advanced ctDNA analyses, including fragmentomics, methylation-based assays such as enzymatic methyl-seq (EM-seq), and circulating histone modification profiling, can capture tumor-specific epigenetic and chromatin features, offering dynamic and complementary insights beyond tissue-based biomarkers.“

4) would revise the following statement, since the data generated in this trial do not suggest that any of the immunotherapy regimens studied here are good enough to be applied in mCRPC - if there is to be a role for immune checkpoint inhibitor therapy in mCRPC beyond pembrolizumab's tumor agnostic indications, it will require better selective biomarkers. other immunotherapy agents, such as bispecific T cell engagers, are showing high response rates in unselected populations of mCRPC patients.

"Overall, these data add to the emerging body of evidence suggesting a potential role for immunotherapy-based regimens in the treatment of select patients with mCRPC"

We agree with the reviewer that better selective biomarkers will be needed to support a role for immune checkpoint inhibitor-based therapies in mCRPC. We believe that the clinical and biomarker data presented in this manuscript represent a step forward in this direction. As such, we have revised the indicated statement to state (page 23): “Overall, these data contribute to our further understanding of a potential role for immunotherapy-based regimens in the treatment of select patients with mCRPC”.