IJC
International Journal of Cancer

The predictive power of tumor mutational burden in lung cancer immunotherapy response is influenced by patients' sex

Shixiang Wang[1,2,3], Jing Zhang[1,2,3], Zaoke He[1,2,3], Kai Wu[4] and Xue-Song Liu[1]

[1]School of Life Science and Technology, ShanghaiTech University, Shanghai, China
[2]Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China
[3]University of Chinese Academy of Sciences, Beijing, China
[4]Department of Thoracic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Immunotherapy, represented by immune checkpoint inhibitors (ICI), is transforming the treatment of cancer. However, only a fraction of patients show response to ICI, and there is an unmet need for biomarkers that will identify patients more likely to respond to ICI. Here we report that the ICI response prediction biomarker tumor mutational burden (TMB) shows significant sex differences. TMB's predictive power is significantly better for female than for male lung cancer patients. Receiver operating characteristic curve analysis was performed and the area under the curve (AUC) was reported to evaluate the predictive power of TMB in lung cancer ICI response. Hazard ratios (HR) of TMB-high *vs.* TMB-low patients were compared between male and female patients. Both AUC and HR differences between female and male are significant in all available independent lung cancer datasets. However, the AUC of programmed death ligand 1 (PD-L1) expression does not show a difference between female and male, suggesting TMB, but not PD-L1 expression has a better predictive power for female than for male lung cancer patients. Our study suggests significant sex differences in the performance of TMB in ICI response prediction. Future development of ICI biomarker should consider sex differences and special efforts should be paid to improve the performance of ICI predictive biomarkers for male lung cancer patients.

Introduction

Immune checkpoint inhibitors (ICI, including anti-PD-1 antibodies, anti-PD-L1 antibodies, anti CTLA-4 antibodies or their combinations) have revolutionized cancer treatment, showing higher efficacy than standard therapies in several cancers, including melanoma,[1–7] nonsmall cell lung cancer (NSCLC),[8–13] head and neck squamous cell carcinoma,[14] renal cell carcinoma (RCC),[15] etc. However, the majority of unselected patients will not respond to ICI. Most tumor types show response rates below 40% to PD-1 inhibition, and the objective response rates of each tumor types are reported to be highly correlated with the tumor mutational burden (TMB) of each tumor types.[16] Multiple factors are reported to affect ICI effectiveness, including: PD-L1 expression,[17,18] TMB,[19,20] DNA mismatch repair deficiency,[21] the degree of cytotoxic T cell infiltration,[22] mutational signature,[23,24] antigen presentation defects,[25,26] interferon signaling,[27] tumor aneuploidy[28] and T-cell gene expression signatures.[29] Currently, efforts are still ongoing to identify robust predictive biomarkers to select patients who would derive the maximum potential benefit from immunotherapies.

Initially, expression of PD-L1 was suggested as a predictive biomarker for ICI response.[30] Pembrolizumab has been approved for NSCLC patients with PD-L1 positive tumors (signal in ≥1% of cells in tumor mass by immunohistochemistry) in second-line therapy or patients with high PD-L1 expression (defined as >50% of cells being PD-L1 positive) in the first line advanced setting, based on two clinical trials.[10,11] However, there are inherent problems for PD-L1 as a predictive biomarker. First, a significant fraction of tumors which are negative for PD-L1 staining still show response to ICI, and many tumors which are positive for PD-L1 do not show response to ICI.[4,31] In addition, PD-L1 is an inducible

Key words: tumor mutational burden, cancer immunotherapy, immune checkpoint inhibitors, lung cancer, sex differences

Abbreviations: AUC: area under the curve; DCB: durable clinical benefit; HR: hazard ratios; ICI: immune checkpoint inhibitors; NDB: no durable benefit.; NSCLC: nonsmall cell lung cancer; ROC: receiver operating characteristic; TCGA: The Cancer Genome Atlas; TMB: tumor mutational burden

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare that they have no competing interests.

Grant sponsor: National Natural Science Foundation of China;
Grant number: 31771373; **Grant sponsor:** ShanghaiTech University;
Grant sponsor: Shanghai Pujiang Program;
Grant number: 16PJ1407400
DOI: 10.1002/ijc.32327
History: Received 28 Oct 2018; Accepted 27 Feb 2019; Online 10 Apr 2019.

Correspondence to: Xue-Song Liu, School of Life Science and Technology, ShanghaiTech University, Shanghai 201203, China, E-mail: liuxs@shanghaitech.edu.cn

What's new?

Only a fraction of cancer patients show response to immune checkpoint inhibitors (ICI), and there remains an unmet need for biomarkers to select patients most likely to benefit. Inherent sex differences in immune response could potentially influence the performance of predictive biomarkers. Here, the authors report that the emerging biomarker tumor mutational burden (TMB) shows significant sex differences, with TMB's predictive power being significantly better for female than for male lung cancer patients. Future development of predictive biomarkers for ICI response should consider sex differences, and the performance of biomarkers for male lung cancer patients should also be improved.

receptor, highly dynamic and the detection and standardization of PD-L1 expression are also difficult.

TMB is an emerging predictive biomarker for ICI response.[16,32] Recent approval of mismatch repair (MMR) deficiency as the first tissue agnostic biomarker for anti-PD-1 response further supports TMB as an independent predictive biomarker, since patients with MMR deficiency also contain high TMB.[21,33] TMB information is usually obtained from whole exome sequencing (WES), and recent development of hybrid capture-based next-generation sequencing (HC-NGS) technique provides additional method for TMB determination.[34] TMB high is associated with improved survival in patients receiving ICI across a wide variety of cancer types, but that there may not be one universal TMB cutpoint associated with improved survival among different cancer types.[35]

Sex and gender difference are known variables that have potential effects on both innate and adaptive immune response.[36] On average, women show stronger immune responses than do men, the severity and prevalence of many infections in women are lower than in men.[36,37] On the other hand, roughly 80% of patients with systemic autoimmune diseases are women.[36,38] The inherent sex differences in immune response could be relevant to the natural course of chronic inflammatory conditions such as cancer. Recently, it has been shown through systematic meta-analysis that male patients derive more benefit in ICI *versus* standard chemotherapy when compared to female patients.[39–41] Overall, for both men and women, immunotherapy was more effective than the control, whether this was a placebo or another type of cancer drug.[39]

Here we hypothesize that the biomarker for predicting immunotherapy response could also be separated based on the patients' sex. We systematically evaluated the performance of current ICI response biomarker TMB and PD-L1 expression using ICI clinical studies with available individual patients' TMB or PD-L1 expression information. Clinical endpoints information include durable clinical benefit (DCB) (partial or stable response lasting >6 months), no durable benefit (NDB; progression of disease ≤6 months of beginning therapy) and progression-free survival (PFS) time.

Materials and Methods

Database search and study design

REMARK guidelines are followed in this report.[42] To analyze TMB and ICI response prediction, we searched PubMed for lung cancer ICI clinical studies with individual patient's TMB information provided (from January 1, 2012, to September 1, 2018). The search terms were ["CTLA-4", "cytotoxic T-lymphocyte-associated protein 4", "PD-1", "programmed death 1", "PD-L1", "Programmed death-ligand 1", "immune checkpoint inhibitor"] AND ["mutational", "mutation", "tumor mutational burden", "tumor mutational load", "genomics" and "genome"] AND ["lung cancer"]. We also reviewed the references of papers included. We excluded studies that enrolled fewer than 10 participants for each sex. In total, 164 studies have been retrieved. Among these studies, only four studies provided detailed TMB number information for individual patients, and those studies that do not provide detailed individual patient's TMB information cannot be selected for further analysis. Among these four studies, one NSCLC dataset[43] only contains 12 samples with available TMB information and is not sufficient for comparing male and female's difference in ICI response. Finally, three NSCLC[20,44,45] datasets are selected for detailed analysis. One NSCLC dataset[44] contains three independent studies, and two of them have also been analyzed independently. To evaluate the predictive power of PD-L1 expression in ICI therapeutic response prediction, only clinical trials with PD-L1 expression data in linear format were included, and in total two datasets[44,45] have been analyzed. Multiple analysis methods were performed to evaluate the predictive power of TMB, including receiver operating characteristic (ROC) curve analysis with DCB/NDB information, hazard ratio (HR) analysis with PFS time. All selected three studies used very similar definition for clinical endpoints. DCB was defined as partial or stable response lasting >6 months, No durable benefit (NDB) was defined as progression of disease ≤6 months of beginning therapy. To evaluate if sex is an independent factor influencing TMB's predictive power, other confounding factors including smoking status, *EGFR* mutation and *ALK* fusion status were also considered in our analysis.

The Cancer Genome Atlas datasets analysis

To analyze the sex differences in mutation load and PD-L1 mRNA expression, we used The Cancer Genome Atlas (TCGA) lung adenocarcinoma datasets,[46] which contain 517 samples with available mutation, mRNA expression and sex status information. For mutation load and PD-L1 protein expression analysis, we used The Cancer Proteome Atlas (TCPA) protein expression database,[47] which contain 260 lung

adenocarcinoma samples with available mutation, PD-L1 protein expression and sex status information.

TMB data normalization

For all three NSCLC ICI clinical studies, TMB values reported in this work were normalized as number of nonsynonymous substitution mutations per million base pairs (MB).

$$\mathrm{TMB} = \frac{\text{Total nonsynonymous substitutions}}{\text{Length of captured region (MB)}}$$

For whole exome sequencing (WES) data reported by Rizvi *et al.*[20] and Hellmann *et al.*,[45] the length of the captured region is 30 (MB). For targeted NGS data reported by Rizvi 2018 dataset,[44] the lengths of the captured region are 0.98, 1.06 and 1.22 (MB) in 341, 410 and 468 gene panels, respectively.

Mutation data download and preprocess

Mutation annotation format (MAF) files of Rizvi 2015 dataset and Rizvi 2018 dataset were downloaded from cBioPortal (http://www.cbioportal.org/). For Hellmann 2018 dataset, we downloaded mutation information file from original article and then used custom R scripts to generate MAF-like file, followed by reannotation using maf2maf.pl script (https://github.com/mskcc/vcf2maf#maf2maf), which is based on Variant Effect Predictor (http://asia.ensembl.org/info/docs/tools/vep/script/index.html). TMB values were then calculated for all three studies. To assure the quality of data extraction for Hellmann 2018 dataset, a correlation analysis was implemented between TMB calculated from reannotated MAF file and the corresponding available nonsynonymous mutation counts from original reference,[45] the Pearson correlation coefficient is more than 0.99.

Receiver operator characteristic curve analysis

The receiver operator characteristic (ROC) curve was generated by plotting the rate of DCB at various threshold settings of TMB. For more detail, the proportion of all DCB patients with TMB above any given cut point (sensitivity) was plotted against the proportion of the no durable benefit (NDB) patients that would also exceed the same cut point (specificity). The area under the curve (AUC) values were reported for each analysis. The significance of the difference between AUC of male and female was determined by unpaired two-tailed Student's *t*-test.

HR analysis

Kaplan–Meier curves of PFS in NSCLC datasets were compared between TMB-high *vs.* TMB-low patients. HR for disease progression or death was reported for each analysis. The pooled HR in male or female were calculated using fixed-effect model. The heterogeneity between male and female HR were evaluated using an interaction test and *p* values were reported. *Q*-test was performed to evaluate the between-study heterogeneity, I^2 were then calculated, which presents the percentage of total variability caused by study heterogeneity.[48] To avoid ecological bias in subgroup analysis, we calculated the within-trial association between TMB's separation effects in HR analysis (TMB-high *vs.* TMB-low) and sex and then pooled these using standard meta-analytical techniques (also called deft approach in Ref. 49).

Statistical analysis

Data between two groups were compared using a two-tailed unpaired Student's *t*-test or Wilcoxon rank-sum test (also known as "Mann–Whitney" test) depending on the normality of data distribution (typically, preprocessed expression data are normally distributed and mutation data show nonnormal distribution). For survival analysis, HRs were calculated by coxph function in R and their 95% CIs were reported, Kaplan–Meier survival curves were modeled by survfit function in R and visualized by survminer package (https://github.com/kassambara/survminer). The log-rank test was used to compare Kaplan–Meier survival curves. All reported *p* values are two-tailed, and for all analyses, $p < 0.05$ is considered statistically significant unless otherwise specified. Statistical analyses were performed using R (version 3.5.0) and Stata 15.0.

Results

Cancer immunotherapy genomics datasets overview

Our database search and literature review identified four lung cancer ICI clinical studies with individual patient's TMB information available. Due to sample size, only three of them are further considered here (Fig. 1). All three studies focused on NSCLC. In total, 349 NSCLC (171 men and 178 women) patients are included in this analysis (Supporting Information Table S1). The majority (240/349 = 69%) of the patients in these clinical trials received monotherapy with anti-PD-(L)1 antibodies, the remaining 109 patients (109/349 = 31%) received anti-PD-(L)1 plus anti-CTLA-4 combination therapy. The TMB information was evaluated by WES or targeted NGS, and no consistently different TMB levels are observed between male and female patients (Supporting Information Table S1). The mutation pattern of these samples are consistent with previously reported NSCLC genomics,[50,51] and no consistent sex differences in the distribution of gene mutations are observed (Supporting Information Fig. S1).

ROC curves of TMB show sex differences in ICI response prediction

To evaluate the predictive power of TMB in ICI response prediction, ROC analysis was performed (Fig. 2*a*). In all three NSCLC datasets analyzed, ROC for TMB in female patients is better than in male patients (Fig. 2*a*). Similarly in the two sub-datasets of Rizvi 2018 study and in all NSCLC combined cases, ROC for TMB in female patients are better than in male patients (Fig. 2*a*). For female and male patients, the average AUC for TMB in NSCLC is 0.755 (95% CI = 0.651–0.859) and 0.554 (95% CI = 0.444–0.664), respectively. The significances

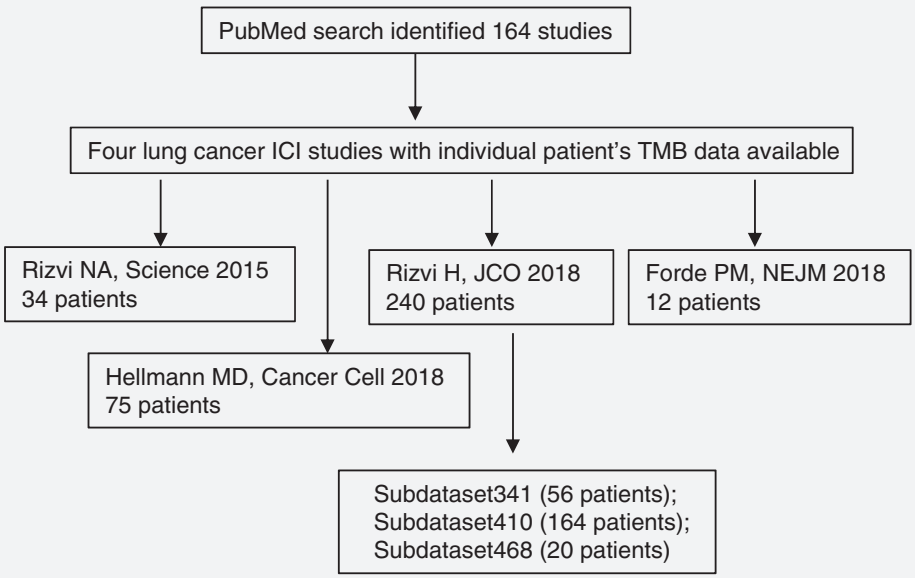

Figure 1. Study selection. We searched PubMed for all lung cancer ICI clinical studies with individual patients' TMB information provided (from January 1, 2012, to September 1, 2018). In total, four studies have been identified. Due to sample size, three of these four studies have been selected for analysis here. Forde PM, NEJM 2018 study does not contain sufficient number of patients for analysis.

of AUC difference are calculated with unpaired two-tailed Student's *t*-test, and reach significance with all NSCLC datasets ($p = 0.043$; Fig. 2*b*). This sex difference in TMB performance is not caused by different smoking status since in both smoker and nonsmoker, similar TMB ROC differences are observed between different sexes (Fig. 3). *EGFR* mutation and *ALK* fusion status also do not influence this sex difference in TMB's performance (Supporting Information Fig. S2). This suggests that sex is an independent factor in determining TMB's performance in ICI prediction.

Prognosis of TMB-high *vs.* TMB-low patients in ICI clinical trials show sex differences

Since the ROC analysis only contains DCB/NDB information of patients, the patients in the same DCB group could have different prognosis as reflected by different overall or PFS time. To further compare the prediction power of TMB in male and female, the HRs of TMB-high *vs.* TMB-low patients was calculated and an across-trial interaction test was performed (Fig. 4*a* and 4*b*). When all three ICI clinical studies are combined, the median value for TMB distribution is about five nonsynonymous mutation/MB (Supporting Information Table S2). Thus, five nonsynonymous mutation/MB was selected as the threshold to separate TMB-high patients from TMB-low patients in our study. In all three NSCLC datasets analyzed, the HR of TMB-high *vs.* TMB-low for female is better than the HR for male patients (Fig. 4*a* and 4*b*). The pooled HR in male and female NSCLC is 0.73 (95%CI 0.41–1.04) and 0.42 (95%CI 0.28–0.56), respectively, and the difference between male and female reaches significance ($p = 0.034$) with an interaction test (Fig. 4*b*), suggesting the separation effect of TMB in NSCLC ICI prognosis prediction is significantly better for female than for male patients. With TMB cutoff varied from 1 to 20, HR of TMB-high *vs.* TMB-low patients in female are always lower than in male, this analysis further supported our conclusion (Supporting Information Fig. S3).

To avoid ecological bias in the across-trial interaction test shown above, a within-trial interaction test was performed.[49] Even more significant difference ($p = 0.002$) was observed between male and female in TMB's separation effect in NSCLC ICI prediction (Fig. 5). This analysis further confirmed the key point of this work that TMB's separation effects in ICI response prediction is significantly influenced by patient's sex. The heterogeneity of sex differences in TMB separation effect (I^2 value = 0 in Fig. 5) is much lower than the heterogeneity of HR difference across different datasets (I^2 value = 74% in Fig. 4*b*). This different heterogeneity implicates that even though each different datasets are heterogeneous; however, the sex differences in TMB's separation effects are strong and consistent in all datasets available. This also explains why the *p* values of within-trial analysis (Fig. 5) is smaller than the *p*-value of across-trial analysis (Fig. 4*b*).

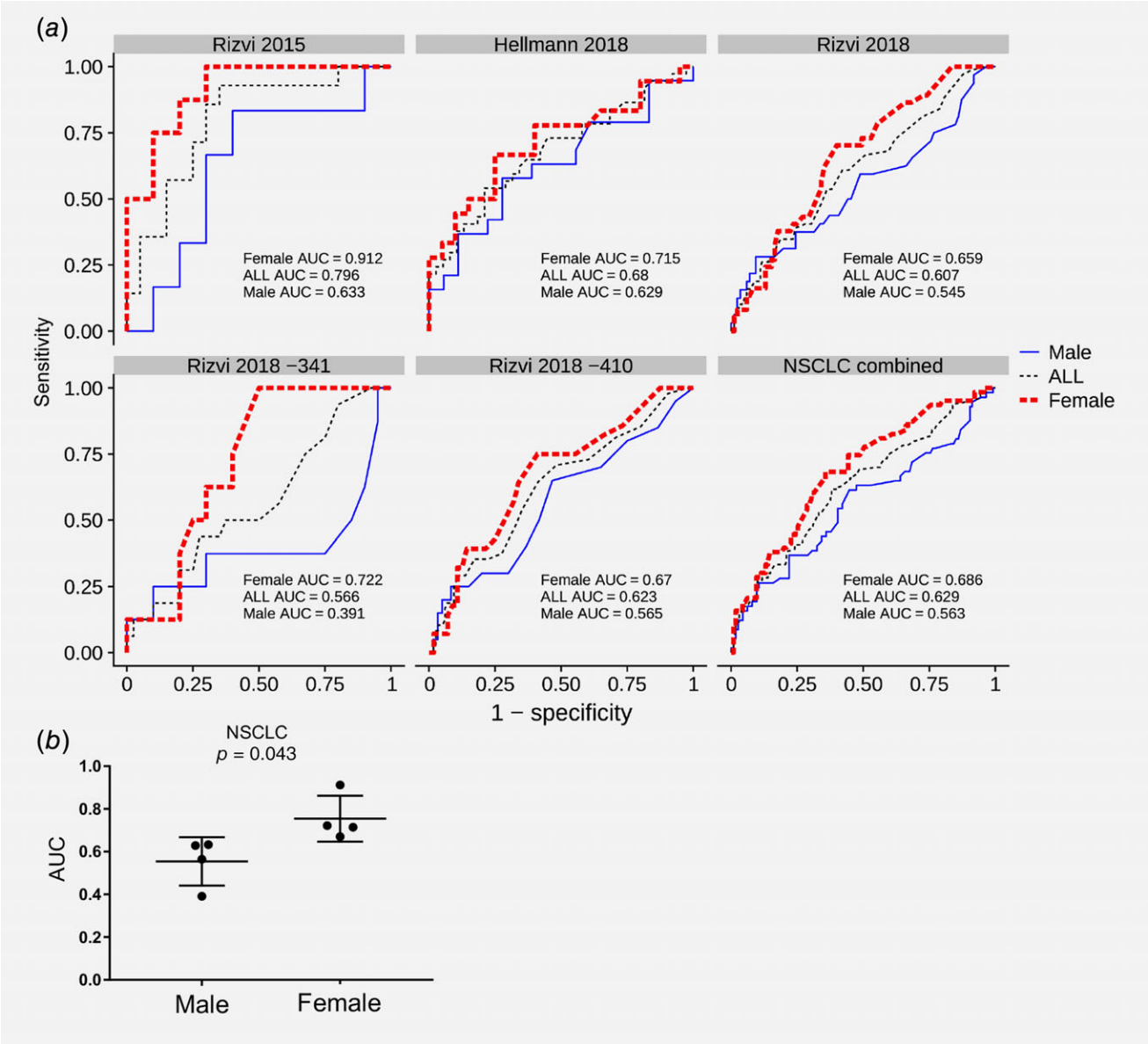

Figure 2. ROC curve analysis for TMB and durable clinical benefit (DCB) prediction in NSCLC datasets. In all three studies, DCB was defined as partial or stable response lasting ›6 months. (*a*) ROC curves for three NSCLC datasets, two subdatasets in Rizvi 2018 publication and combined NSCLC datasets. (*b*) Area under the curve (AUC) was statistically compared between male and female NSCLC datasets using two-tailed unpaired Student's *t*-test. [Color figure can be viewed at wileyonlinelibrary.com]

TMB differences between DCB and NDB reach significance in female, but not in male NSCLC patients

We further compared the TMB levels in NSCLC patients with DCB *vs.* patients with NDB and observed that the differences reach significance only in female patients but not in male patients (Fig. 6), and this TMB difference between female and male is consistent among all available NSCLC datasets (Fig. 6). This observation is consistent with the ROC analysis. The AUC of TMB for male NSCLC is around 0.6, suggesting a relatively poor prediction effects, as a consequence, the TMB does not show a significant difference between male DCB and male NDB patients.

Next, the correlation between ICI clinical effects (DCB, NDB) and TMB status was evaluated with Fisher's exact test. Significant correlations are only observed for female patients (Supporting Information Table S3). This result is also consistent with the above ROC and HR analysis.

PD-L1 expression does not show sex differences in ICI response prediction

Since PD-L1 expression is an independent biomarker for predicting ICI response, we compared the predictive power of PD-L1 expression between male and female patients. In

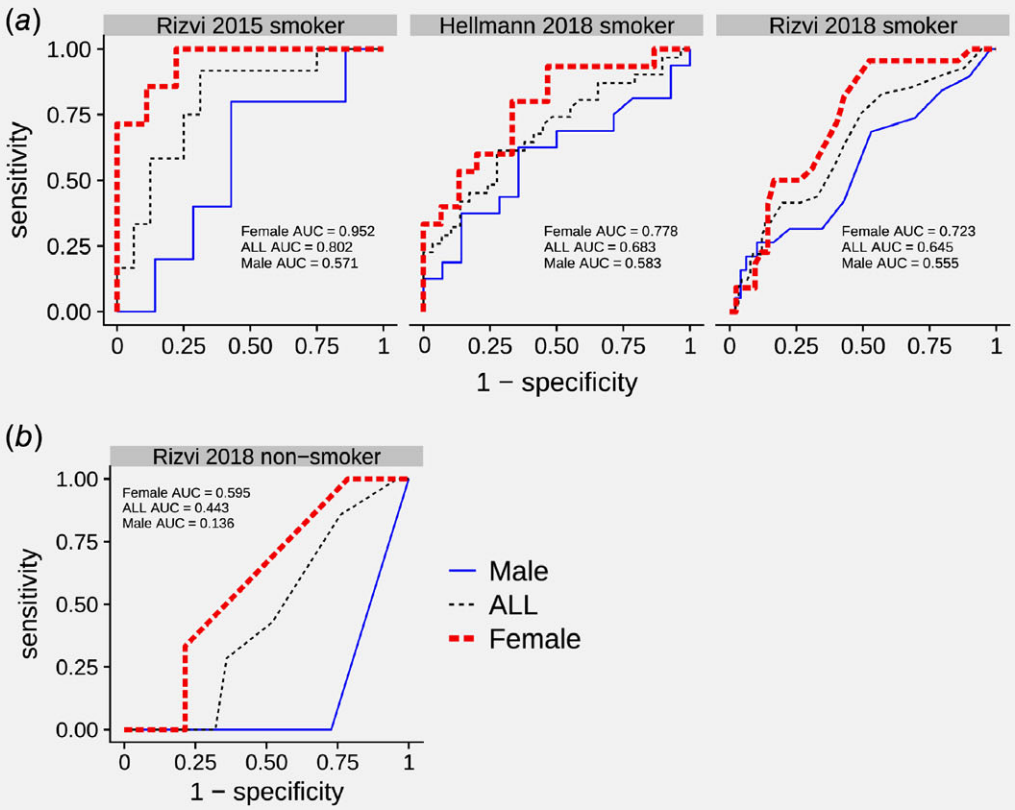

Figure 3. Effects of smoking status on ROC difference between female and male NSCLC patients. (*a*) ROC curves of patients with smoking history from three NSCLC datasets. (*b*) ROC curve of patients without smoking history in Rizvi 2018 dataset, the other two datasets do not contain sufficient number of nonsmoker patients for this ROC analysis. [Color figure can be viewed at wileyonlinelibrary.com]

general, the predictive power of PD-L1, as reflected by the AUC value, is worse than TMB. No difference is observed between male and female with PD-L1 expression as ICI prognosis biomarker (Supporting Information Fig. S4).

Sex differences in TMB and tumor immunoediting

The TMB performance difference between male and female will have immediate clinical implications in patient stratification. We speculate the inherent differences between female and male in immune response could account for this observed TMB performance difference. Tumors with high TMB are usually more immunogenic compared to tumors with low TMB due to the increased presentation of neoantigens on tumor cell surface.[52] Strong immune response can eliminate highly immunogenic tumors, however, in the case of immunoediting, highly immunogenic tumors could evade host immune response through immune checkpoint regulation.[53] It has already been established that women usually show stronger immune response than men.[36,37] Without immunoediting, the chance for TMB-high tumors be eliminated in female patients is bigger than in male patients. Thus, we predict that female patients with high TMB should have stronger immune suppression signaling than male. The expression of PD-L1 could be served as an indicator for these immune suppression signaling. As expected, with an independent TCGA lung cancer genomics database, we found that significantly elevated PD-L1 mRNA and protein expression are specifically observed in female patients with high TMB but not male patients (Supporting Information Fig. S5). Since PD-L1 expression is one indicator for the strength of immune suppression signaling, and PD-L1 expression itself is directly involved in ICI function. As expected, the predictive power of PD-L1 expression is not affected by sex differences.

Discussion

There is an unmet need for biomarkers that will identify patients more likely to respond to ICI. Recent studies suggest TMB as an efficient pan-cancer biomarker for ICI response.[16,32] Here we performed systematic analysis to

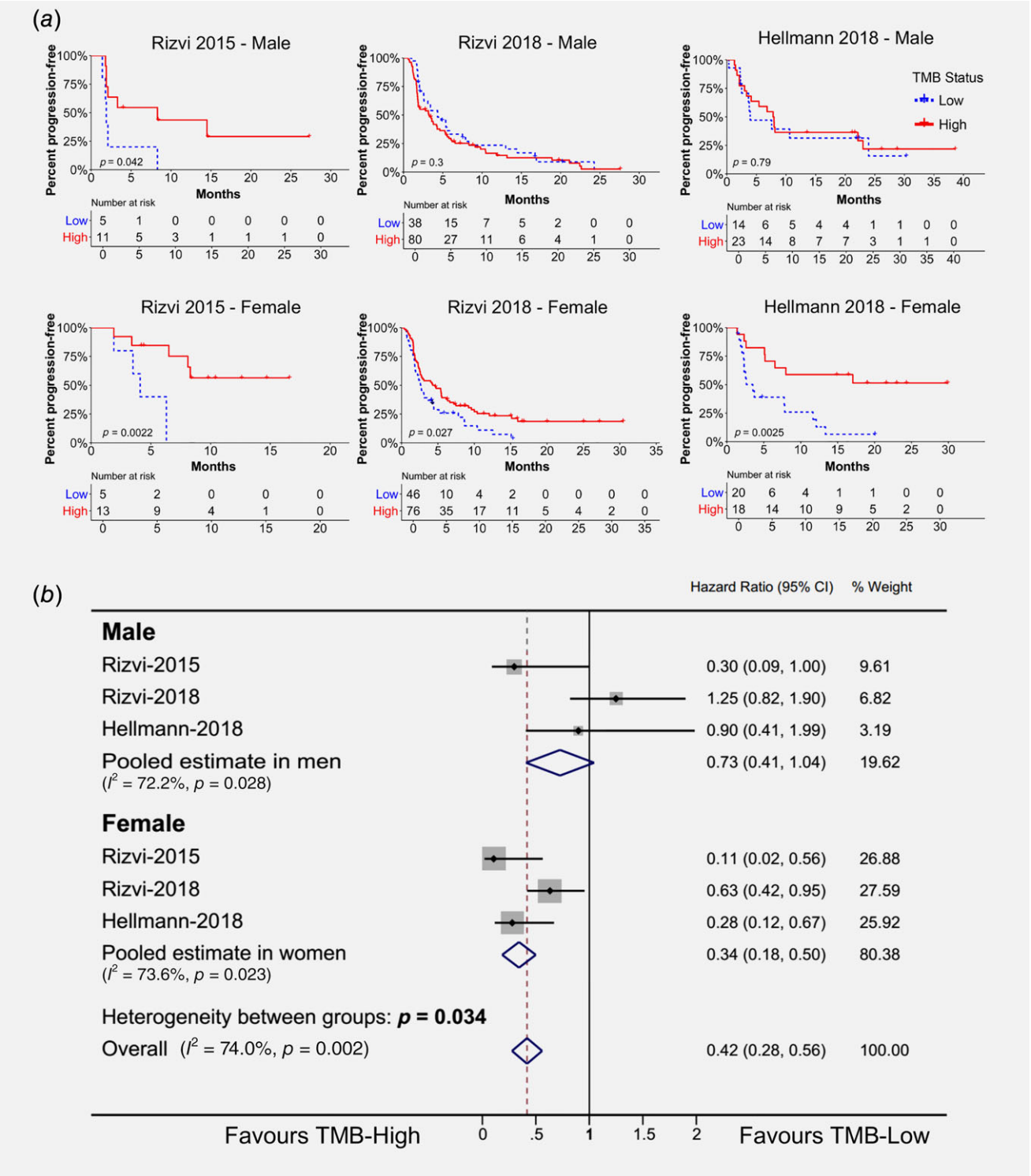

Figure 4. Kaplan–Meier (KM) progression-free survival (PFS) curves and HR analysis based on TMB level in female and male NSCLC patients. (*a*) KM curves of patients with high or low level of TMB in three NSCLC datasets are shown. The threshold to separate TMB-high from TMB-low was set as five nonsynonymous mutations per MB based on TMB's distribution pattern. Patients with TMB ›5 nonsynonymous mutations per MB was defined as TMB-high, and patients with TMB ‹5 nonsynonymous mutations per MB was defined as TMB-low. (*b*) HR for disease progression analysis comparing TMB-high with TMB-low. Squares represent study-specific HR. Horizontal lines indicate 95% CI. Diamonds indicate the pooled HR, calculated separately in females and males, with their corresponding 95% CI. The *p* value (0.034) represents heterogeneity by patients' sex. [Color figure can be viewed at wileyonlinelibrary.com]

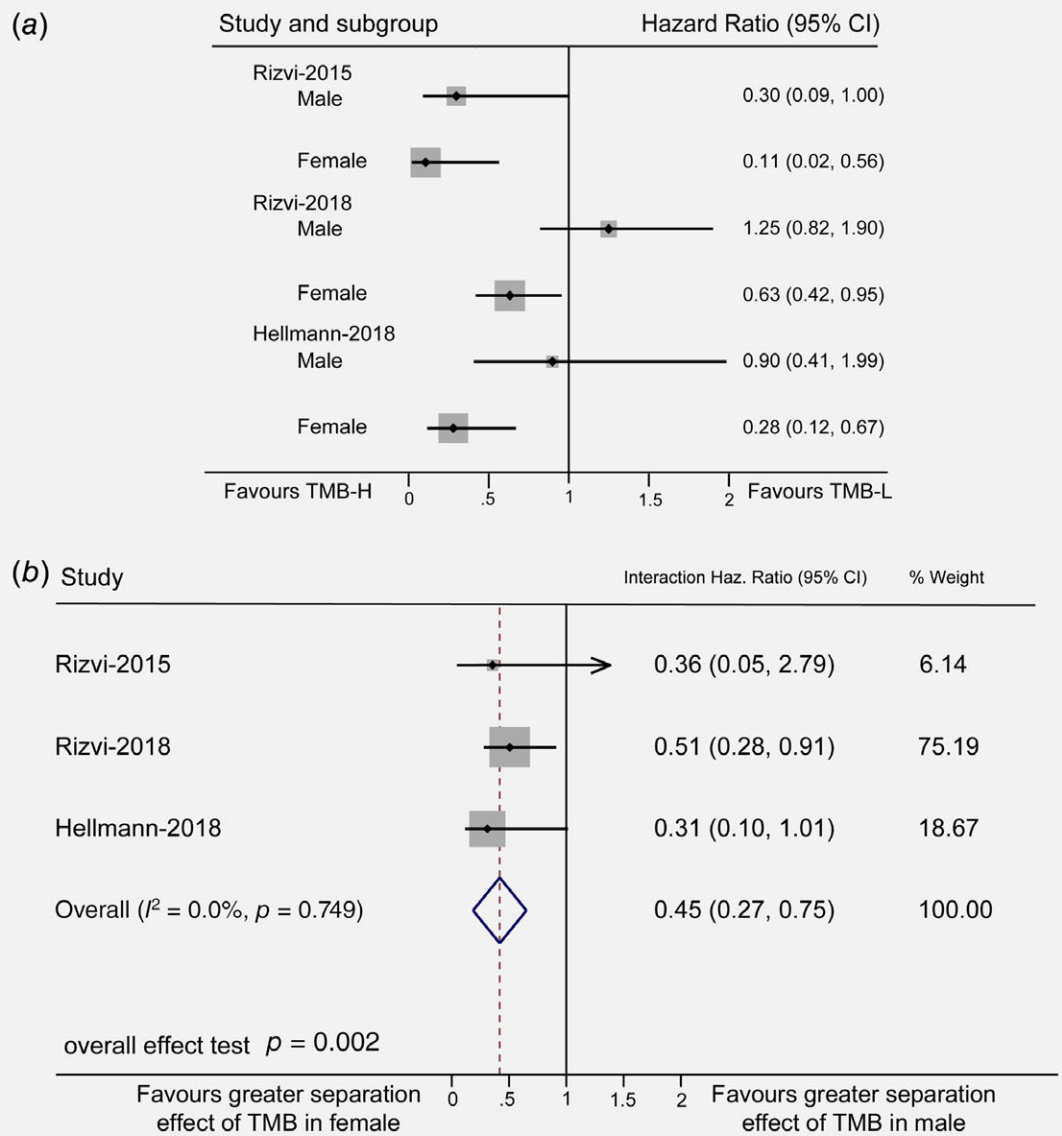

Figure 5. Hazard ratio (HR) for disease progression analysis comparing TMB-high with TMB-low in the three NSCLC datasets as in Figure 4. In the top panel, squares represent study-specific HR. Horizontal lines indicate 95% CI. In the bottom panel, interaction between sex differences and TMB's separation effects in HR analysis are calculated within each three datasets. Squares represent study-specific interactions and horizontal lines indicate the 95% CI, Diamonds indicate the pooled interactions. The overall heterogeneity by patients' sex was evaluated using interaction test and *p*-value (0.002) is reported. [Color figure can be viewed at wileyonlinelibrary.com]

compare the performance of ICI prediction biomarkers between male and female. Generally, TMB show improved performance compared to PD-L1 in ICI response prediction. We suggest for the first time that for female patients the performance of TMB, but not PD-L1 expression in ICI response prediction is significantly better than for male NSCLC patients.

With different statistical analysis methods, the difference between female and male in TMB's predictive power in ICI response reach significance, even with a very limited number of datasets available for our study. In all three NSCLC datasets and two subdatasets of Rizvi 2018 dataset, consistent differences between male and female in TMB performance are observed. Currently, we can only obtain four independent lung cancer

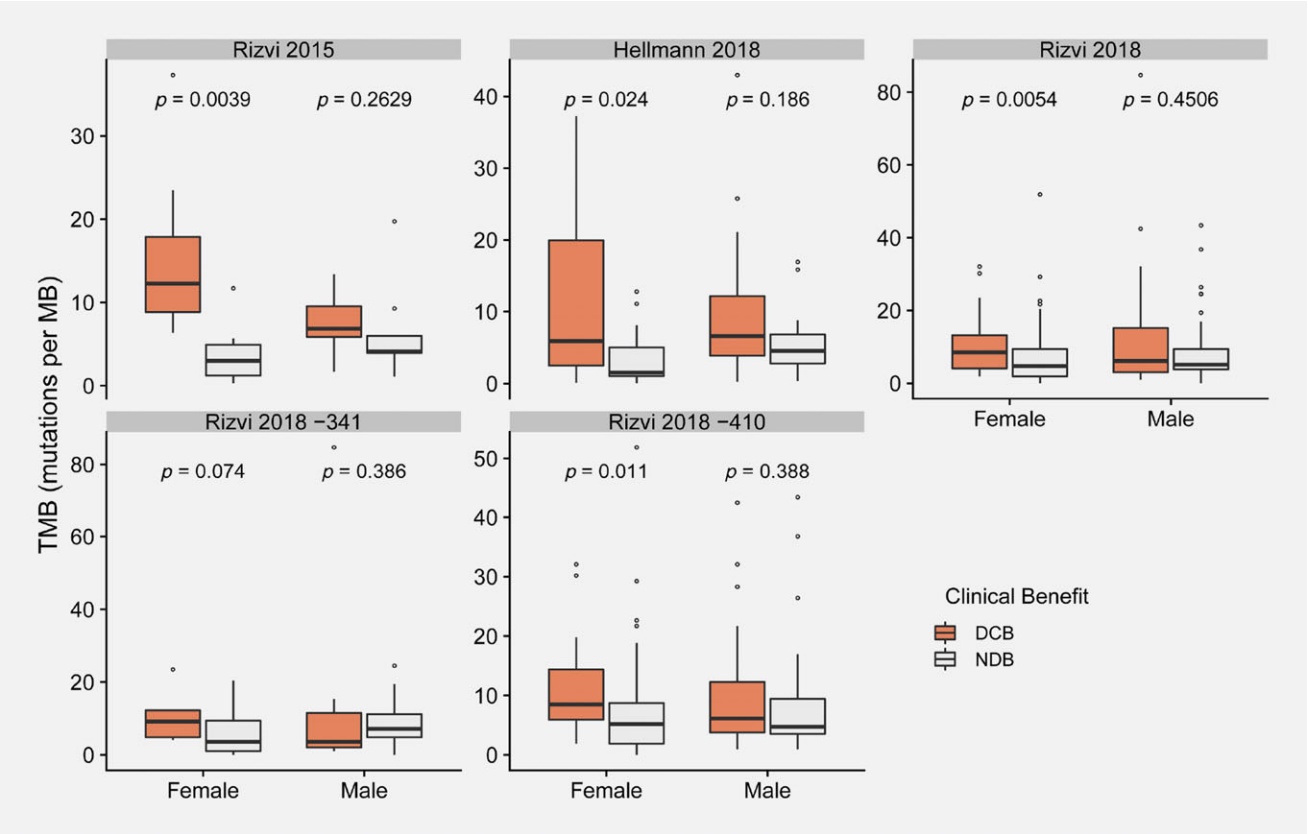

Figure 6. TMB difference between patients with DCB and patients with NDB only reach significance in female but not male NSCLC patients. For three NSCLC ICI datasets and two subdatasets of Rizvi 2018, the TMB level was compared between DCB and NDB group in female and male patients. The boxplot is bounded by the first and third quartile with a horizontal line at the median. Two-tailed Mann–Whitney test was performed and *p* values are shown. [Color figure can be viewed at wileyonlinelibrary.com]

genomic datasets with ICI clinical information. Even though we got a statistically significant conclusion based on this limited number of datasets, more datasets are needed in the future to further validate this.

PD-L1 expression is another independent biomarker for predicting ICI response, even though its performance is problematic.[31] Our study suggests that the predictive power of PD-L1 expression on ICI response is not affected by the patients' sex. Probably because PD-L1 expression itself is directly involved in ICI function, and consequently, the predictive power of PD-L1 expression is not affected by sex differences.

Recent studies demonstrated that males derive more benefit in ICI *vs.* chemotherapy than females.[39–41] Here, we report that the ICI predictive biomarker TMB has a better separation effect in female than in male patients. These two findings actually do not contradict each other. We speculate that the inherent strong immune response in female patients could account for both of these two observations. Based on this study, TMB is less likely to separate male responder from nonresponder in ICI therapy. In the future, novel biomarkers or biomarker combinations should be considered especially for male lung cancer patients.

Our study has the following immediate clinical implications. The future design of ICI predictive biomarker, especially TMB based biomarker should consider sex differences. The performance of TMB in male ICI response prediction is relatively poor, thus we need to exert more effort on the improvement of predictive biomarker specifically for male patients.

Acknowledgements

We thank the authors and participating patients of these publications (Rizvi NA, *et al.* Science 2015; Rizvi H, *et al.* J Clin Oncol, 2018; Hellmann MD, *et al.* Cancer cell 2018) for providing the data for this analysis. Thank TCGA project for making lung cancer genomics data available for analysis. Thank David Fisher of MRC Clinical Trials Unit at UCL for critical help in Meta-analysis. We thank Raymond Shuter for editing the text. Thank ShanghaiTech University High-Performance Computing Public Service Platform for computing services. Thank members of Liu lab for helpful discussion. This work was supported in part by the Shanghai Pujiang Program (16PJ1407400), The National Natural Science Foundation of China (31771373) and startup funding from ShanghaiTech University.

Author contributions

Data collection and statistical analysis: Wang S and Zhang J; Analysis and interpretation of data: Wang S, He Z and Liu X-S; Critical discussion: Wu K; Study design, supervision and article writing: Liu X-S.

References

1. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010; 363:711–23.
2. Robert C, Thomas L, Bondarenko I, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med* 2011;364:2517–26.
3. Ribas A, Kefford R, Marshall MA, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *J Clin Oncol* 2013;31:616–22.
4. Robert C, Long GV, Brady B, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* 2015;372:320–30.
5. Robert C, Schachter J, Long GV, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N Engl J Med* 2015;372:2521–32.
6. Hodi FS, Chesney J, Pavlick AC, et al. Combined nivolumab and ipilimumab versus ipilimumab alone in patients with advanced melanoma: 2-year overall survival outcomes in a multicentre, randomised, controlled, phase 2 trial. *Lancet Oncol* 2016;17:1558–68.
7. Larkin J, Minor D, D'Angelo S, et al. Overall survival in patients with advanced melanoma who received Nivolumab versus Investigator's choice chemotherapy in CheckMate 037: a randomized, controlled, open-label phase III trial. *J Clin Oncol* 2018;36:383–90.
8. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 2015;373: 1627–39.
9. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med* 2015;373: 123–35.
10. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* 2016;387:1540–50.
11. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med* 2016;375:1823–33.
12. Carbone DP, Reck M, Paz-Ares L, et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med* 2017;376:2415–26.
13. Govindan R, Szczesna A, Ahn MJ, et al. Phase III trial of Ipilimumab combined with paclitaxel and carboplatin in advanced squamous non-small-cell lung cancer. *J Clin Oncol* 2017;35:3449–57.
14. Ferris RL, Blumenschein G Jr, Fayette J, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med* 2016; 375:1856–67.
15. Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus Everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015;373:1803–13.
16. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med* 2017;377:2500–1.
17. Herbst RS, Soria JC, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 2014;515: 563–7.
18. Shukuya T, Carbone DP. Predictive markers for the efficacy of anti-PD-1/PD-L1 antibodies in lung cancer. *J Thorac Oncol* 2016;11:976–88.
19. Snyder A, Makarov V, Merghoub T, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 2014;371:2189–99.
20. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.
21. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;372:2509–20.
22. Tang H, Wang Y, Chlewicki LK, et al. Facilitating T cell infiltration in tumor microenvironment overcomes resistance to PD-L1 blockade. *Cancer Cell* 2016;29:285–96.
23. Wang S, Jia M, He Z, et al. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. *Oncogene* 2018;37: 3924–36.
24. Miao D, Margolis CA, Vokes NI, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet* 2018;50:1271–81.
25. Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med* 2016;375: 819–29.
26. Chowell D, Morris LGT, Grigg CM, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* 2018;359:582–7.
27. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127: 2930–40.
28. Davoli T, Uno H, Wooten EC, et al. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* 2017;355:eaaf8399.
29. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* 2018;24:1550–8.
30. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol* 2010;28:3167–75.
31. Fusi A, Festino L, Botti G, et al. PD-L1 expression as a potential predictive biomarker. *Lancet Oncol* 2015;16:1285–7.
32. Goodman AM, Kato S, Bazhenova L, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* 2017;16:2598–608.
33. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;357:409–13.
34. Campesato LF, Barroso-Sousa R, Jimenez L, et al. Comprehensive cancer-gene panels can be used to estimate mutational load and predict clinical benefit to PD-1 blockade in clinical practice. *Oncotarget* 2015;6:34221–7.
35. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* 2019;51:202–6.
36. Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol* 2016;16:626–38.
37. vom Steeg LG, Klein SL. SeXX matters in infectious disease pathogenesis. *PLoS Pathog* 2016;12: e1005374.
38. Whitacre CC, Reingold SC, O'Looney PA. A gender gap in autoimmunity. *Science* 1999;283:1277–8.
39. Conforti F, Pala L, Bagnardi V, et al. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. *Lancet Oncol* 2018;19:737–46.
40. Botticelli A, Onesti CE, Zizzari I, et al. The sexist behaviour of immune checkpoint inhibitors in cancer therapy? *Oncotarget* 2017;8:99336–46.
41. Pinto JA, Vallejos CS, Raez LE, et al. Gender and outcomes in non-small cell lung cancer: an old prognostic variable comes back for targeted therapy and immunotherapy? *ESMO Open* 2018;3:e000344.
42. Altman DG, McShane LM, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLoS Med* 2012;9:e1001216.
43. Forde PM, Chaft JE, Smith KN, et al. Neoadjuvant PD-1 blockade in resectable lung cancer. *N Engl J Med* 2018;378:1976–86.
44. Rizvi H, Sanchez-Vega F, La K, et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. *J Clin Oncol* 2018;36:633–41.
45. Hellmann MD, Nathanson T, Rizvi H, et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. *Cancer Cell* 2018;33: 843–52 e4.
46. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The cancer genome atlas Pan-cancer analysis project. *Nat Genet* 2013;45:1113–20.
47. Li J, Lu Y, Akbani R, et al. TCPA: a resource for cancer functional proteomics data. *Nat Methods* 2013;10:1046–7.
48. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327:557–60.
49. Fisher DJ, Carpenter JR, Morris TP, et al. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;356:j573.
50. Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* 2012;150: 1121–34.
51. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014;511:543–50.
52. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science* 2015;348:69–74.
53. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* 2011; 331:1565–70.