JAMA Oncology | **Original Investigation**

Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade
A Systematic Review and Meta-analysis

Steve Lu; Julie E. Stein, MD; David L. Rimm, MD, PhD; Daphne W. Wang, MS; J. Michael Bell;
Douglas B. Johnson, MD; Jeffrey A. Sosman, MD; Kurt A. Schalper, MD, PhD; Robert A. Anders, MD, PhD;
Hao Wang, PhD; Clifford Hoyt, MS; Drew M. Pardoll, MD, PhD; Ludmila Danilova, PhD; Janis M. Taube, MD

+ **Supplemental content**

IMPORTANCE PD-L1 (programmed cell death ligand 1) immunohistochemistry (IHC), tumor mutational burden (TMB), gene expression profiling (GEP), and multiplex immunohistochemistry/immunofluorescence (mIHC/IF) assays have been used to assess pretreatment tumor tissue to predict response to anti–PD-1/PD-L1 therapies. However, the relative diagnostic performance of these modalities has yet to be established.

OBJECTIVE To compare studies that assessed the diagnostic accuracy of PD-L1 IHC, TMB, GEP, and mIHC/IF in predicting response to anti–PD-1/PD-L1 therapy.

EVIDENCE REVIEW A search of PubMed (from inception to June 2018) and 2013 to 2018 annual meeting abstracts from the American Association for Cancer Research, American Society of Clinical Oncology, European Society for Medical Oncology, and Society for Immunotherapy of Cancer was conducted to identify studies that examined the use of PD-L1 IHC, TMB, GEP, and mIHC/IF assays to determine objective response to anti–PD-1/PD-L1 therapy. For PD-L1 IHC, only clinical trials that resulted in US Food and Drug Administration approval of indications for anti–PD-1/PD-L1 were included. Studies combining more than 1 modality were also included. Preferred Reporting Items for Systematic Reviews and Meta-analysis guidelines were followed. Two reviewers independently extracted the clinical outcomes and test results for each individual study.

MAIN OUTCOMES AND MEASURES Summary receiver operating characteristic (sROC) curves; their associated area under the curve (AUC); and pooled sensitivity, specificity, positive and negative predictive values (PPV, NPV), and positive and negative likelihood ratios (LR+ and LR−) for each assay modality.

RESULTS Tumor specimens representing over 10 different solid tumor types in 8135 patients were assayed, and the results were correlated with anti–PD-1/PD-L1 response. When each modality was evaluated with sROC curves, mIHC/IF had a significantly higher AUC (0.79) compared with PD-L1 IHC (AUC, 0.65, $P < .001$), GEP (AUC, 0.65, $P = .003$), and TMB (AUC, 0.69, $P = .049$). When multiple different modalities were combined such as PD-L1 IHC and/or GEP + TMB, the AUC drew nearer to that of mIHC/IF (0.74). All modalities demonstrated comparable NPV and LR−, whereas mIHC/IF demonstrated higher PPV (0.63) and LR+ (2.86) than the other approaches.

CONCLUSIONS AND RELEVANCE In this meta-analysis, tumor mutational burden, PD-L1 IHC, and GEP demonstrated comparable AUCs in predicting response to anti–PD-1/PD-L1 treatment. Multiplex immunohistochemistry/IF and multimodality biomarker strategies appear to be associated with improved performance over PD-L1 IHC, TMB, or GEP alone. Further studies with mIHC/IF and composite approaches with a larger number of patients will be required to confirm these findings. Additional study is also required to determine the most predictive analyte combinations and to determine whether biomarker modality performance varies by tumor type.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Janis M. Taube, MD, Division of Dermatopathology Johns Hopkins University, 600 N Wolfe St, Blalock Building Room 907, Baltimore, MD 21287 (jtaube1@jhmi.edu).

S ubstantial effort is ongoing to identify predictors of response and resistance to anti-PD-1/PD-L1 (anti-programmed cell death 1/programmed cell death ligand 1) immunotherapy. Expression of PD-L1 protein was the first candidate biomarker associated with response to anti-PD-1 therapy.[1,2] Multiple PD-L1 immunohistochemistry (IHC) assays have been approved by the US Food and Drug Administration (FDA) as a companion or complementary diagnostic for patients with non–small cell lung cancer (NSCLC), melanoma, bladder cancer, gastric cancer, and cervical carcinoma to help preselect patients for anti-PD-1/PD-L1 therapy.[3,4] These IHC assays have been used to stratify patients in clinical trials. Positive PD-L1 IHC has also recently been shown to enrich for response to combination therapy with anti-PD-1/cytotoxic T lymphocyte antigen-4 in patients with NSCLC.[5] However, although PD-L1 expression has been shown to correlate with response to therapy in certain tumor types, the association is not absolute. As a diagnostic assay, PD-L1 IHC has several limitations: multiple different assays are available, the predictive significance of tumor cell vs immune cell expression varies by tumor type,[6] the scoring of immune cell PD-L1 expression by pathologists has poor interobserver reproducibility,[7] and the PD-L1 expression is commonly reduced to a digital readout (+ vs −) without assessing its expression in the greater context of the tumor microenvironment (TME) (eg, association with immune cells suggests an adaptive pattern of expression).[8,9]

More recent biomarker approaches include assessing tumor mutational burden (TMB), gene expression profiling (GEP), and quantifying multiple proteins using multiplex IHC/immunofluorescence (mIHC/IF). The assessment of TMB as a biomarker for response sensitivity to immunotherapy is predicated on the concept that more mutations yield more T cell–recognized tumor neoantigens, potentially resulting in stronger antitumor immune responses when the PD-1 checkpoint is blocked.[10] Tumor mutational burden was first shown to be associated with response to cytotoxic T lymphocyte antigen-4 blockade in patients with melanoma,[11] followed by patients with NSCLC treated with anti-PD-1 therapy (pembrolizumab),[12] and now extends to numerous solid tumor types.[13]

Gene expression profiling allows for the simultaneous assessment of a number of parameters. The mRNA transcript levels of inflammatory genes, immune checkpoint genes, and even oncogenes have been included in various gene panels. This approach has a continuous output and has been used to develop "response signatures" for several tumor types, most of which include an interferon (IFN) gamma gene signature as a major pillar of the assay.[14] However, this approach lacks information on cellular coexpression and geography within the TME. In contrast, mIHC/IF allows for the simultaneous visualization of multiple IHC/IF protein markers in situ on the same tissue section. As such, it provides a spatial component that can be used to generate cell density metrics for a given tissue region or to assess the distance between 2 given cell types. Coexpression of multiple markers on a single cell can also be readily visualized. Similar to PD-L1 IHC, TMB, and GEP, mIHC/IF assays have now been associated with response to anti-PD-1/PD-L1 therapies in multiple different tumor types.[9,15-17]

Key Points

Question What is the relative diagnostic accuracy of different biomarker assay modalities in predicting clinical response to anti-PD-1/PD-L1 (programmed cell death 1/programmed cell death ligand 1) therapy?

Findings In this systematic review and meta-analysis involving tumor specimens from 8135 patients, multiplex immunohistochemistry/immunofluorescence (mIHC/IF) had significantly higher diagnostic accuracy than PD-L1 IHC, tumor mutational burden, or gene expression profiling in predicting clinical response to anti-PD-1/PD-L1 therapy and was similar to multimodality cross-platform composite approaches, such as PD-L1 IHC + tumor mutational burden.

Meaning Multiplex immunohistochemistry/IF facilitates quantification of protein coexpression on immune cell subsets and assessment of their spatial arrangements; initial findings suggest that mIHC/IF has diagnostic accuracy comparable to multimodality cross-platform composite approaches in predicting response to anti-PD-1/PD-L1.

The purpose of the current analysis was to compare existing data on the diagnostic accuracy of PD-L1 IHC, TMB, GEP, and mIHC/IF biomarker modalities in predicting response to anti-PD-1/PD-L1 therapy. We performed a meta-analysis using summary receiver operating characteristic (sROC) curves to determine the relative area under the curves (AUCs) as a global metric of each approach's ability to discriminate between responders and nonresponders to therapy. We also determined the relative sensitivity, specificity, predictive values, and likelihood ratios of these emerging approaches and compared them with PD-L1 IHC and each other.

Methods

This meta-analysis was conducted in adherence to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) statement.[18] Two independent reviewers (S.L. and J.E.S.) performed the literature search, assessed eligibility criteria, and performed data extraction (S.L. and D.W.W.).

Search Strategy
The DailyMed website[19] was used to identify clinical trials that tested PD-L1 status for an association with response to therapy using IHC and were cited in association with FDA-approved indications for nivolumab, pembrolizumab, atezolizumab, durvalumab, and avelumab monotherapies. The national clinical trial number was gathered and PubMed was searched to identify the associated clinical trial report.

For all other assay modalities, predefined search criteria were used to conduct electronic searches of PubMed (from inception to June 2018). Searches were limited to human studies with English translation available. The search syntax included the following terms: (mutational burden OR mutational load OR mutational density OR mutational landscape OR genomic landscape OR whole exome sequencing OR gene

expression profiling OR gene signature OR mRNA OR multiplex immunofluorescence OR multiplex immunohistochemistry OR spatial profiling) AND (anti–PD-1 OR anti–PD-L1 OR nivolumab OR pembrolizumab OR atezolizumab OR durvalumab OR avelumab OR BMS-936558 OR BMS-936559 OR MK-3475 OR MPDL3280A OR MEDI4736 OR MSB0010718C). In addition, the 2013 to 2018 annual meeting abstracts from the American Society of Clinical Oncology, the European Society for Medical Oncology, the Society for Immunotherapy of Cancer, and the American Association for Cancer Research were searched using the keywords "nivolumab," "pembrolizumab," "atezolizumab," "durvalumab," and "avelumab." A second search was performed in November 2018 for the definitive manuscript related to abstracts identified during the June 2018 search, and the extracted data was updated to reflect the final publication. Several experts in the field were also surveyed to determine if there were any additional publications or conference abstracts that were not revealed by the electronic search.

Eligibility Criteria

Studies that correlated pretreatment tissue-based biomarkers of interest with objective response rate (ORR) (ie, complete response and partial response), progression free survival (PFS), or overall survival (OS) were included if they had at least 15 patients treated with anti–PD-1/PD-L1 monotherapy. Tumor mutational burden, GEP, and mIHC/IF studies that enrolled patients based on PD-L1 status were excluded from all of the single-modality assay approaches but were included in the multimodality group if they met all other criteria. Studies involving patients with hematologic cancers and flow cytometry studies on tumor lysates were also excluded. When studies were identified with overlapping participants, the study with the largest tested population was included.

Data Extraction and Calculation of Diagnostic Accuracy Measures

The following data categories were extracted from the included studies: study name, national clinical trial number, therapy received, biomarkers tested, number of patients tested with each biomarker, and the year of publication or conference presentation. Biomarker results (+ vs −) as they related to objective response to therapy were extracted for each study and used to generate a 2 by 2 contingency table showing true-positive, false-positive, true-negative, and false-negative test results for each study. These values were used to calculate sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR−).

In 7 of 56 analyses, ORR was not provided, and OS was gathered.[17,20-23] In 2 instances, OS was not available, and PFS was used.[24,25] When using OS or PFS, sensitivity and specificity values were calculated at 6 months, 1 year, and 2 years. The Youden statistic (a measure of informedness) was calculated to determine the optimal time point for each individual study that maximized sensitivity and specificity. This point was then included in the meta-analysis to allow the best possible representation for each study. For PD-L1 IHC studies, if

multiple thresholds of assay positivity were presented, the FDA-approved cutoff point for the associated companion/complimentary diagnostic for that indication was used. For 1 study in which multiple cutoff points were presented, there was no accompanying PD-L1 IHC diagnostic for that indication. In that instance, the Youden statistic was also performed, and the best performing threshold was chosen.

Statistical Analysis

The calculated sensitivity and specificity from each individual biomarker analysis were plotted, and a curve was fit to the points using both weighted and unweighted linear regression models. For the weighted approach, the hierarchical DerSimonian and Laird method was used; the Moses-Littenberg model was applied for the unweighted approach. The sROC curves and the resultant AUC were used to measure the association between the different assay modalities and ORR. The AUCs were compared for statistically significant differences using the Hanley and McNeil method.[26] When TMB raw data were available, the upper tertile of TMB vs the lower 2 tertiles within a given study were evaluated against response.[24,27]

For each modality, the measures of diagnostic accuracy were pooled to generate an overall metric allowing for comparisons between each modality (eg, pooled sensitivity for PD-L1 IHC vs pooled sensitivity for GEP). Because methodological heterogeneity between included studies was anticipated, a random-effects model was used for pooling the analyzed parameters. Possible publication bias was assessed by examining a funnel plot of the effect size for each study against the reciprocal of its standard error (metaphor package in R[28]).

The dedicated meta-analysis software Meta-DiSc was used for evaluation of the various biomarker tests.[29] All statistical tests were 2-sided and $P < .05$ was considered statistically significant. A Bonferroni correction was applied to account for multiple comparisons.

Results

Search Results

The search strategy identified 45 eligible reports that assayed either PD-L1 IHC (n = 24),[20,30-51] TMB (n = 10),[12,13,16,17,21,22,52-56] GEP (n = 9),[13,21-23,40,54,57-59] mIHC/IF (n = 7),[9,15-17,25,60,61] or multimodality (PD-L1 IHC + TMB and/or GEP, n = 6)[13,24,52,59] and correlated the results with response to anti–PD-1/PD-L1 therapy (**Figure 1** and eTable 1 in the Supplement). Eleven studies reported either multiple different clinical study cohorts or more than 1 of the individual biomarker approaches, resulting in a total of 56 individual analyses. eTable 2 in the Supplement provides a summary of study size, median patient age, sex, tumor stage, tumor type, treatment, ORR, and assay performance characteristics for each trial. Clinical covariates such as age, sex, tumor stage, and tumor type did not vary between assay modalities (eFigure 1 in the Supplement). In total, specimens from 8135 patients with more than 10 different solid tumor types are represented in the meta-analysis.

Figure 1. PRISMA Flowchart
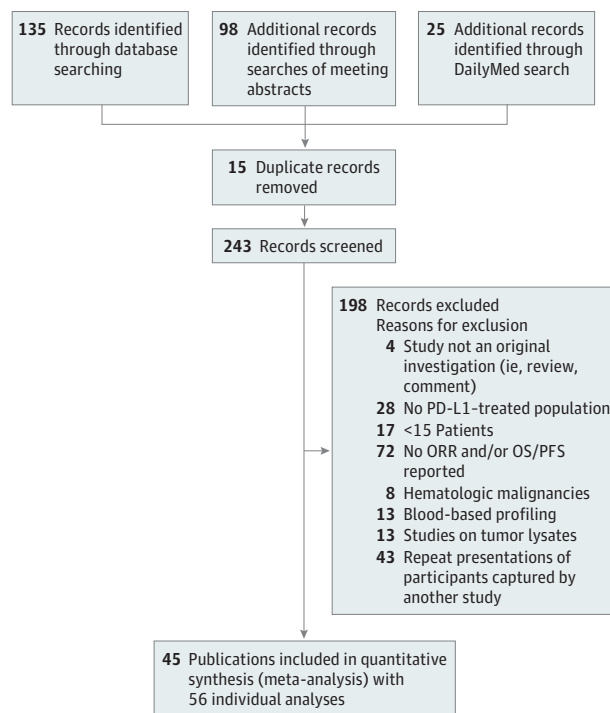
eTable 1 in the Supplement lists the included studies. PD-L1 indicates programmed cell death ligand 1; PFS, progression-free survival; ORR, objective response rate; OS, overall survival.

sROC Curves

The derived sensitivity and 1-specificity values from each report were plotted, **Figure** 2A. The points were fitted with sROC curves using 2 different approaches: 1 that weighted each study by the number of patients enrolled, and 1 in which each study was given equal weight. By weighted sROC curves, mIHC/IF had a significantly higher AUC (0.79) than PD-L1 IHC (0.65, $P < .001$), GEP (0.65, $P = .003$), and TMB (0.69, $P = .049$) (Figure 2B). Neither TMB nor GEP had an AUC profile that significantly differed from that of PD-L1 IHC. Most of the GEP studies included an IFN gamma gene signature. However, the broad category of GEP as reported herein also includes reports of other gene signatures associated with therapeutic resistance (Nos. 35 and 42 in eTable 1 in the Supplement). When those reports were excluded, the sROC curve AUC for only IFN gamma-based GEP studies was 0.65. When each reported study was given equal weight, irrespective of the number of patients enrolled, the relative AUC results remained consistent (Figure 2C). Similarly, when the 9 (16%) of 56 analyses[17,20-25] with ORR imputed from OS/PFS data were excluded from the analysis, the AUC results remained consistent (eFigure 2 in the Supplement).

Approaches that combined variables across multiple platforms were also explored. The sROC curves from the multimodality biomarkers approached the AUC of mIHC/IF (0.74 vs 0.79, $P = .48$) (**Figure 3**). This supports earlier findings showing that measures of an "inflamed" TME combined with TMB have additive prognostic or PV over either parameter alone.[13,62]

Analyses of Stable Disease, Tumor Type, and TMB Thresholds

Of 47 studies that reported ORR, 45 (96%) included stable disease (SD), with progressive disease (PD) reported under the heading of "nonresponders."[9,12,13,15,30-61] Of the 18 studies that reported SD as a separate category that could be reanalyzed,[9,12,36,38-41,44,47,49,51,52,55,56,58,61] we compared the sROC curves for each modality when SD was included with responders rather than with PD, and the resulting AUCs did not noticeably differ (eFigure 3 in the Supplement). Subgroup analysis by tumor type was also performed (eFigure 4 in the Supplement), but a larger number of studies will be necessary to resolve any potential differences in assay modality performance by tumor type.

Subgroup analysis was also performed on the study cohorts that assessed for TMB to explore the influence of various approaches used for thresholding cases into high TMB vs low TMB categories. The original analysis was performed using the threshold of positivity set by each individual study. For the 6 of 10 studies that had raw data available, we used a uniform thresholding approach, whereby the upper tertile of TMB was considered a positive test and the remainder was considered a negative test.[12,17,52,54-56] In this subset of cases, the profiles of the sROC curves for TMB were not noticeably affected by the use of a uniform approach, with AUCs of 0.71 in both analyses (eFigure 5 in the Supplement).

Additionally, although studies focused on microsatellite instability (MSI) were not included in this meta-analysis, a study reporting MSI as a predictor of response to anti–PD-1/PD-L1 therapy[63] is included alongside the mIHC/IF data points and sROC curves for comparison (eFigure 6 in the Supplement) and demonstrates that mIHC/IF assays may have sensitivity and specificity similar to that of MSI status when predicting response to anti-PD-1/PD-L1 therapy. Similarly, a study that included a machine learning component for biomarker discovery using TMB and GEP data[64] is shown in eFigure 6 in the Supplement for interest. The high sensitivity and specificity achieved in that study suggests that machine learning algorithms may further improve diagnostic accuracy.

PVs and LRs

The PPV and NPV for each study are plotted in **Figure 4**A and eFigure 7 in the Supplement. Most modalities provide relatively high NPV. PD-L1 immunohistochemistry, GEP, and multimodality biomarkers have relatively low PPV. In contrast, mIHC/IF assays consistently demonstrate a high PPV (**Table**). The PPV and NPV plots for NSCLC, melanoma, and urothelial carcinoma are shown separately in eFigure 4 in the Supplement.

The pooled LRs for each modality are shown in Figure 4B and eTable 3 in the Supplement. Multiplex immunohistochemistry/IF and the multimodality biomarkers have LRs+ that are significantly higher than those of PD-L1 IHC (2.86 and 2.76, respectively, vs 1.51), indicating that they are less likely to produce a false-positive result. The pooled LRs– for GEP (0.65), TMB (0.62), PD-L1 IHC (0.69), and multimodality approaches (0.60) were essentially equivalent. The mIHC/IF studies demonstrated a trend toward an improved LR– compared with the

Figure 2. Summary Receiver Operating Characteristic Curve Analysis by Assay Modality for Responders vs Nonresponders
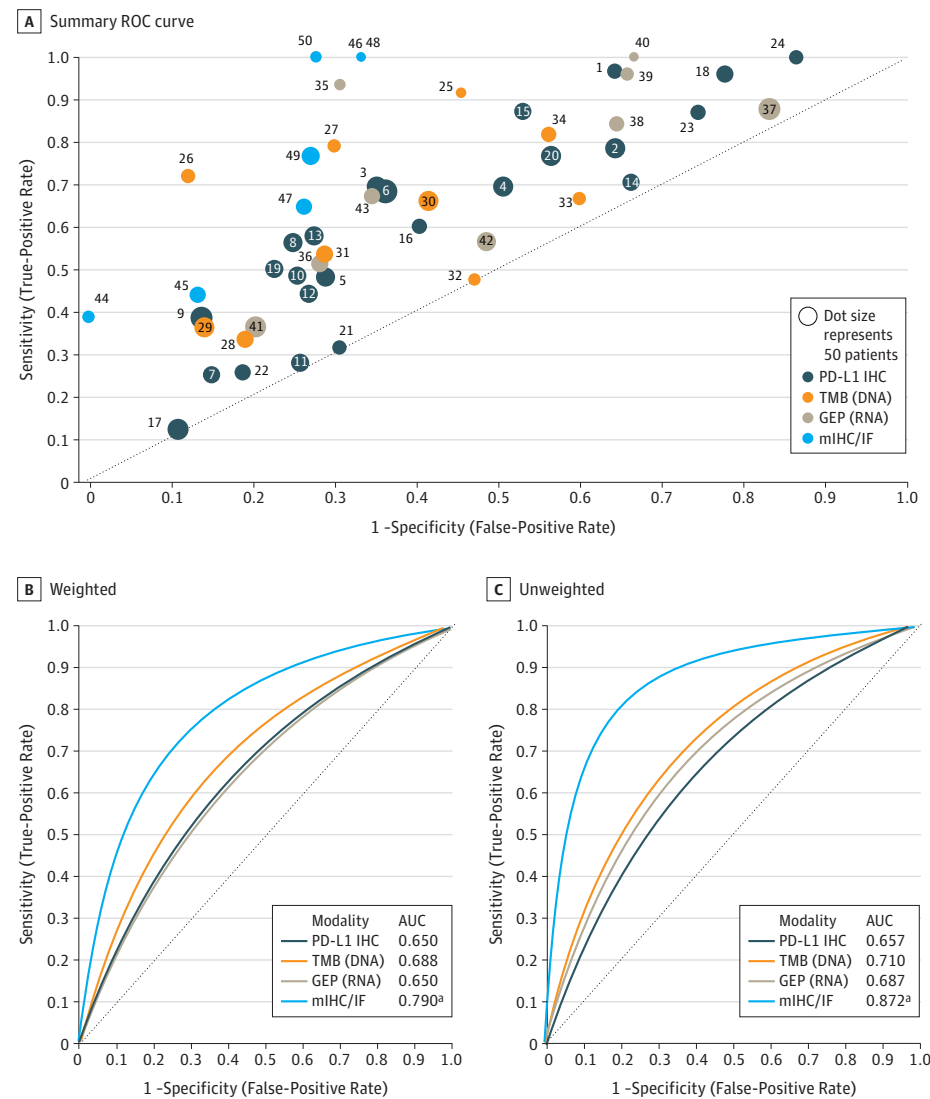
A, Fifty-six analyses from the primary literature correlating different biomarker modalities with patient responses after anti–PD-1/PD-L1 therapy were analyzed. The sensitivity and 1–specificity of the assay for each individual publication is shown by a single dot (number on the dot correlates with reference list in eTable 1 in the Supplement). The size of each dot is proportionate to the size of the studied cohort. Linear regression models weighted (B) by the number of patients in each study and unweighted (C) (ie, each study treated equally) were used to generate summary receiver operating characteristic [sROC] curves for each assay modality). The multiplex immunohistochemistry/ immunofluorescence (mIHC/IF) has a significantly higher area under the curve (AUC) than PD-L1 (programmed cell death ligand 1) IHC, tumor mutational burden (TMB), and gene expression profiling (GEP) by weighted approach and PD-L1 IHC and TMB by unweighted approach.

ª Indicates statistical significance (*P* < .05), Hanley and McNeil method.

Pooled Sensitivity and Specificity

The pooled sensitivity and specificity for each modality are summarized in the Table. Gene expression profiling and mIHC/IF have improved sensitivity compared with that of PD-L1 IHC, whereas TMB, mIHC/IF, and multimodality have improved specificity. Gene expression profiling actually has a lower specificity for response than PD-L1 IHC. It is not immediately clear why this is the case, especially because many of the GEP panels detect PD-L1 RNA. Possible explanations include a lack of correlation between RNA and protein levels for the markers of interest or the fact that RNA is less stable than protein and thus is harder to preserve and later detect as an analyte. In addition, protein-based measurements are commonly associated with a higher dynamic range than RNA analytes.[65]

other modalities, although this difference was not statistically significant.

Publication Bias

Because there are relatively fewer publications with small cohort sizes for the newer modalities compared with those for PD-L1 IHC, a funnel plot was used to assess potential publication bias. Bias appeared to be trivial, with no effect on major findings (data available from the authors).[66]

Discussion

Pretreatment predictive biomarkers for immuno-oncology are sought after to help match individual patients to the treatment regimen most likely to be of benefit. Predictive biomarkers may also accelerate clinical trials and FDA approvals, aid in cost containment, and help members of the medical community provide accurate patient guidance. Currently, a number of different biomarker modalities are being pursued. The most common tissue-based biomarker

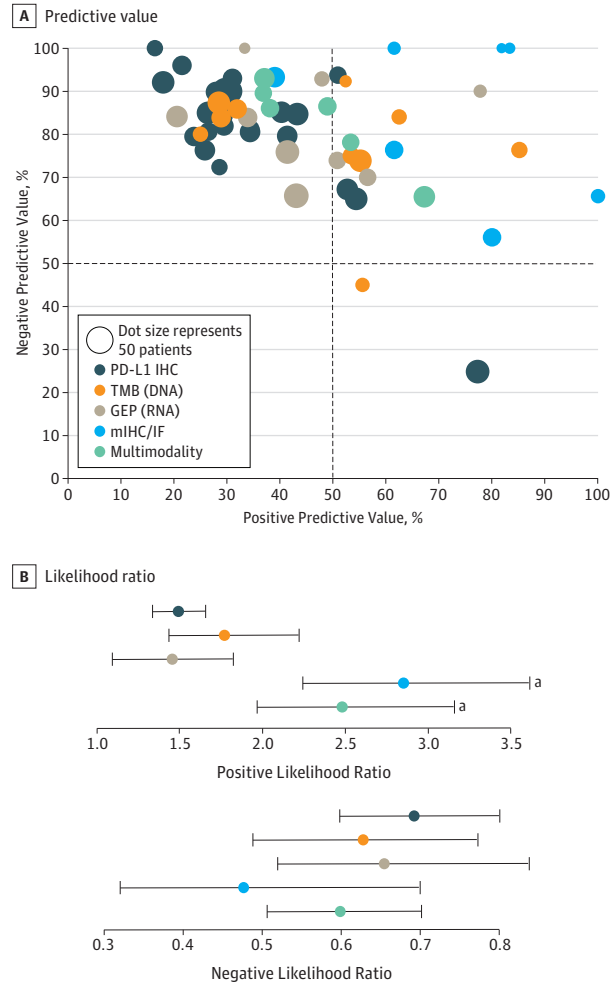## Figure 3. Summary Receiver Operating Characteristic (sROC) Curve Analysis
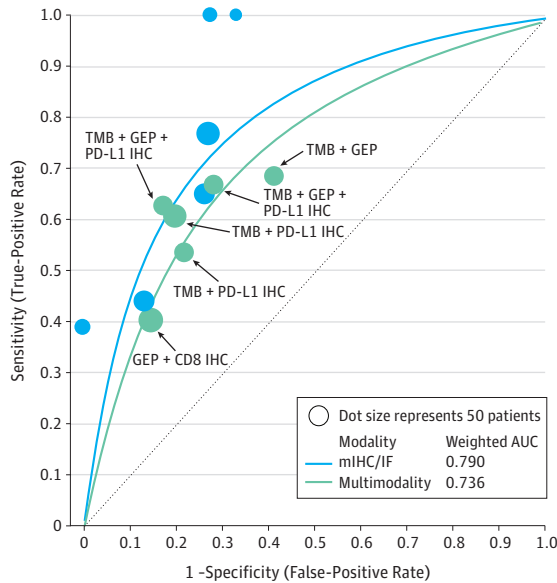
The Figure shows that multimodality biomarkers have an sROC curve comparable to that of multiplex immunohistochemistry/immunofluorescence. AUC indicates area under the curve; GEP, gene expression profiling; IHC, immunohistochemistry; mIHC/IF, multiplex immunohistochemistry/ immunofluorescence; PD-L1, programmed cell death ligand 1; TMB, tumor mutational burden.

Figure 4. Predictive Values for Each Individual Study and Pooled Likelihood Ratios by Biomarker Assay Modality

A, Dots in the upper left quadrant represent studies that reported high negative predictive values and were good at excluding patients from anti–PD-1/PD-L1 treatment. Dots in the upper right quadrant represent studies that reported both high negative and positive predictive values, meaning they are suitable for excluding patients who should not be treated and for selecting patients who will respond. Multiplex immunohistochemistry/immunofluorescence (mIHC/IF) can help both rule in and rule out response to anti–PD-1/PD-L1 therapy. The study number correlating to the individual dots is provided in eFigure 7 in the Supplement. B, Multiplex immunohistochemistry/IF has a better likelihood ratio (LR−) than other tested biomarker approaches, whereas both mIHC/IF and multimodality approaches have significantly higher LRs(+) (eTable 3 in the Supplement).

[a] Statistically significant.

approaches for predicting response to anti–PD-1/PD-L1 therapy include PD-L1 IHC, TMB, GEP, and mIHC/IF. This meta-analysis provides an overview of the relative diagnostic accuracy of these approaches. We found that mIHC/IF has a significantly higher AUC than PD-L1 IHC, GEP, and TMB and has improved LR+ and LR− compared with the other biomarker approaches. PD-L1 IHC, the most well-established biomarker for response to anti–PD-1/PD-L1 therapy, has one of the lowest AUCs and has generally poor LRs.

The AUC serves as a global measure of how well an assay can distinguish between 2 groups—in this instance—responders and nonresponders to anti–PD-1/PD-L1 therapy. Consensus guidelines regarding acceptable AUCs for companion and complimentary diagnostic tests do not exist. However, some authors have suggested that diagnostic tests used for patient selection should have AUCs of 0.80 or higher.[67-69] In this study, we demonstrate that mIHC/IF has an AUC in this range and that combining different biomarker modalities approaches can also result in an AUC that approaches this target threshold.

It is possible that 2 biomarker approaches could demonstrate near equivalent AUCs, yet have differing sensitivities and specificities. Additional measures of diagnostic accuracy are useful for further assessing how modalities perform at ruling in vs ruling out a patient for a given therapy. Most approaches had a similar ability to identify patients who were not likely to respond to therapy, as indicated by comparable NPVs and LRs−. The notable difference was in identifying patients most likely to respond to therapy. In this regard, mIHC/IF had a significantly higher PPV and LR+. This means that there are fewer

false-positive tests, that is, patients who would be treated but would not respond to therapy. This translates practically to being able to match a patient to a potentially more effective treatment sooner as well as not exposing a patient to potential immune-related adverse effects if they are less likely to respond. Furthermore, because an average treatment course typically costs more than $120 000, using biomarker strategies with improved diagnostic accuracy may help avoid considerable costs to the health care system for much less likely anticipated benefit.

Table. Comparison of Pooled Diagnostic Sensitivity and Specificity and Predictive Values for Responders vs Nonresponders After Anti–PD-1/PD-L1 Therapy Between Different Assay Modalities[a]

Modality	Pooled Sensitivity	Pooled Specificity	Pooled PPVs	Pooled NPVs
PD-L1 IHC (n = 24)	0.50 (0.48-0.53)	0.63 (0.62-0.65)	0.34 (0.32-0.36)	0.78 (0.76-0.79)
TMB (n = 10)	0.57 (0.51-0.62)	0.70 (0.66-0.73)	0.42 (0.38-0.47)	0.80 (0.77-0.83)
GEP (n = 9)	0.71 (0.67-0.75)	0.51 (0.48-0.54)	0.42 (0.39-0.46)	0.77 (0.74-0.81)
mIHC/IF (n = 7)	0.60 (0.53-0.66)	0.78 (0.73-0.82)	0.63 (0.56-0.70)	0.75 (0.70-0.80)
Multimodality (n = 6)	0.58 (0.50-0.65)	0.79 (0.75-0.82)	0.41 (0.33-0.48)	0.88 (0.85-0.91)

Abbreviations: GEP, gene expression profiling; mIHC/IF, multiplex immunohistochemistry/immunofluorescence; PD-L1 IHC, programmed cell death ligand 1 immunohistochemistry; TMB, tumor mutational burden.

[a] All data are reported as a proportion (95% CI). Nonoverlapping 95% CIs suggest statistical significance.

The improved performance of mIHC/IF cannot simply be attributed to a larger number of analytes studied because the GEP methods assayed a median of 17 analytes vs 2 or 3 for mIHC/IF. Instead, these findings suggest improved diagnostic benefit when spatial relationships and protein coexpression on specific cellular subsets are assessed. The studies of mIHC/IF assays included herein evaluated parameters such as PD-1 to PD-L1 proximity (n = 2 studies[15,16]), CD8+ cell density within a specific TME compartment (eg, intratumoral/peritumoral defined by IHC/IF tumor marker, n = 3[9,60,61]), or coexpressed markers indicating T-cell activation (n = 2[17,25]). Spatial and coexpression assessments are not represented in GEP assays where specimens were homogenized before being assayed. There are emerging high multiplex protein detection approaches, including multispectral platforms,[15,70,71] serial stain and strip,[72,73] imaging mass cytometry,[74,75] and multiplexed ion beam imaging.[76,77] Although it is anticipated that the inclusion of additional markers will yield improved PV, large data sets for such comparisons are not yet available.

Many of the modalities test for parameters representing an "inflamed" TME. The aforementioned mIHC/IF assays characterize T-cell activation states, immune checkpoint expression, and/or the density of specific T-cell subsets. Many GEP studies contain an IFN gamma gene signature, and PD-L1 IHC expression may be driven by IFN gamma following T-cell infiltration.[8,78] In contrast, TMB does not necessarily correlate with an "inflamed" TME[62,79] but represents the possibility of there being immunogenic mutation-associated neoantigens present within the tumor. An analysis performed by Danilova et al[62] on melanoma specimens from The Cancer Genome Atlas data set showed that TMB and an inflamed phenotype were distinct variables when predicting patient survival and that TMB specifically had added prognostic value when tumors were less inflamed. The concept that TMB and a marker of an inflamed TME have potential additive value is supported in the present study by the fact that combined biomarker approaches such as PD-L1 IHC + TMB demonstrated a higher AUC and LR+ compared with that of PD-L1 IHC, TMB, or GEP alone. Importantly, the combination of mIHC/IF + TMB has yet to be tested to determine whether the incorporation of TMB can further elevate the mIHC/IF AUC.

Limitations

The limitations of this study include the fact that there were different assays represented within a single-assay type category. For example, the studies of PD-L1 IHC included herein used different companion diagnostic assays with different PD-L1 antibodies, thresholds of positivity, and scoring systems. Similarly, the mIHC/IF assays tested different protein targets using different platforms as described earlier, and the number of detected mRNA targets ranged from 1 (IFN gamma) to 26 different gene sets representing about 690 individual genes. In contrast, TMB arguably represents a more uniform output. Furthermore, the patient outcome tested in this study was ORR. Although it is recognized that ORR correlates with survival,[80] it would be of great interest to specifically test for an association between these assay modalities and PFS and/or OS; however, such long-term follow-up is not yet available for most studies. Finally, it is a significant limitation that mIHC/IF is the newest method of those tested. Even though the analysis was weighted by specimen numbers, the mIHC/IF test numbers used in the study were less than 10% of PD-L1 test numbers, thus warranting caution in the interpretation of this modality. However, given the early promise of mIHC/IF, we look forward to broader future testing.

Conclusions

The increasing number of published studies reporting on predictive biomarkers for response to anti–PD-1/PD-L1 therapy has provided an opportunity to assess the accuracy of individual assay modalities. This meta-analysis suggests that mIHC/IF merits further investigation. The relative success of mIHC/IF in predicting patient response also provides insight into the spatial importance of tumor-immune interactions and the contribution of protein marker coexpression. Importantly, this difference was achieved with a relatively "low plex," that is, an average of 2 to 3 markers examined. Future improvements in diagnostic accuracy are likely to be made by increasing the number of markers detected in the mIHC/IF format and by developing multiplex, multimodal approaches that potentially combine GEP and/or TMB with mIHC/IF.

Author Affiliations: Department of Dermatology, Johns Hopkins Medical Institutions, Baltimore, Maryland (Lu, Stein, D. W. Wang, Bell, Taube); Department of Pathology, Yale University School of Medicine, New Haven, Connecticut (Rimm, Schalper); Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee (Johnson); Department of Medicine, Division of Hematology-Oncology, Northwestern University Medical Center, and Robert H. Lurie Cancer Center, Chicago, Illinois (Sosman); Department of Pathology, Johns Hopkins Medical Institutions,

Baltimore, Maryland (Anders, Taube); Division of Biostatistics & Bioinformatics at the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins Medical Institutions, Baltimore, Maryland (H. Wang, Danilova); Akoya Biosciences, Hopkinton, Massachusetts (Hoyt); Department of Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland (Pardoll, Taube); Bloomberg–Kimmel Institute for Cancer Immunotherapy, Johns Hopkins Medical Institutions, Baltimore, Maryland (Pardoll, Danilova, Taube).

REFERENCES

1. Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med*. 2012;366(26):2443-2454. doi:10.1056/NEJMoa1200690

2. Taube JM, Klein A, Brahmer JR, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res*. 2014;20(19):5064-5074. doi:10.1158/1078-0432.CCR-13-3271

3. Taube JM, Galon J, Sholl LM, et al. Implications of the tumor immune microenvironment for staging and therapeutics. *Mod Pathol*. 2018;31(2):214-234. doi:10.1038/modpathol.2017.156

4. Gong J, Chehrazi-Raffle A, Reddi S, Salgia R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. *J Immunother Cancer*. 2018;6(1):8. doi:10.1186/s40425-018-0316-z

5. Hellmann MD, Rizvi NA, Goldman JW, et al. Nivolumab plus ipilimumab as first-line treatment for advanced non-small-cell lung cancer (CheckMate 012): results of an open-label, phase 1, multicohort study. *Lancet Oncol*. 2017;18(1):31-41. doi:10.1016/S1470-2045(16)30624-6

6. Kulangara K, Zhang N, Corigliano E, et al. Clinical utility of the combined positive score for programmed death ligand-1 expression and the approval of pembrolizumab for treatment of gastric cancer. *Arch Pathol Lab Med*. 2019;143(3):330-337. doi:10.5858/arpa.2018-0043-OA

7. Rimm DL, Han G, Taube JM, et al. A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non-small cell lung cancer. *JAMA Oncol*. 2017;3(8):1051-1058. doi:10.1001/jamaoncol.2017.0013

8. Taube JM, Anders RA, Young GD, et al. Colocalization of inflammatory response with B7-h1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape. *Sci Transl Med*. 2012;4(127):127ra37. doi:10.1126/scitranslmed.3003689

9. Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. 2014;515(7528):568-571. doi:10.1038/nature13954

10. Gubin MM, Schreiber RD. Cancer. the odds of immunotherapy success. *Science*. 2015;350(6257):158-159. doi:10.1126/science.aad4140

11. 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(23):2189-2199. doi:10.1056/NEJMoa1406498

12. 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(6230):124-128. doi:10.1126/science.aaa1348

13. Cristescu R, Mogg R, Ayers M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science*. 2018;362(6411):eaar3593. doi:10.1126/science.aar3593

14. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest*. 2017;127(8):2930-2940. doi:10.1172/JCI91190

15. Giraldo NA, Nguyen P, Engle EL, et al. Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma and association with response to pembrolizumab. *J Immunother Cancer*. 2018;6(1):99. doi:10.1186/s40425-018-0404-0

16. Johnson DB, Bordeaux J, Kim JY, et al. Quantitative spatial profiling of PD-1/PD-L1 interaction and HLA-DR/IDO-1 predicts improved outcomes of anti-PD-1 therapies in metastatic melanoma. *Clin Cancer Res*. 2018;24(21):5250-5260.

17. Gettinger SN, Choi J, Mani N, et al. A dormant TIL phenotype defines non-small cell lung carcinomas sensitive to immune checkpoint blockers. *Nat Commun*. 2018;9(1):3196. doi:10.1038/s41467-018-05032-8

18. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097

19. National Institutes of Health. US National Library of Medicine. DailyMed website. https://dailymed.nlm.nih.gov/dailymed/. Accessed May 10, 2019.

20. Motzer RJ, Escudier B, McDermott DF, et al; CheckMate 025 Investigators. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. 2015;373(19):1803-1813. doi:10.1056/NEJMoa1510665

21. Mariathasan S, Turley SJ, Nickles D, et al. TGFβ attenuates tumour response to PD-L1 blockade by

contributing to exclusion of T cells. *Nature*. 2018; 554(7693):544-548. doi:10.1038/nature25501

22. Powles T, Durán I, van der Heijden MS, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet*. 2018;391(10122):748-757. doi:10.1016/S0140-6736(17)33297-X

23. Fehrenbacher L, Spira A, Ballinger M, et al; POPLAR Study Group. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837-1846. doi:10.1016/S0140-6736(16)00587-0

24. Carbone DP, Reck M, Paz-Ares L, et al; CheckMate 026 Investigators. First-line nivolumab in stage IV or recurrent non–small-cell lung cancer. *N Engl J Med*. 2017;376(25):2415-2426. doi:10.1056/NEJMoa1613493

25. Wong PF, Smithy JW, Blenman KR, Kluger HM, Rimm DL. Quantitative assessment of tumor-infiltrating lymphocytes and immunotherapy outcome in metastatic melanoma [Abstract 3638]. *Cancer Res*. 2018;78(13)(suppl):3638. doi:10.1158/1538-7445.AM2018-3638

26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983;148(3):839-843. doi:10.1148/radiology.148.3.6878708

27. Hellmann MD, Callahan MK, Awad MM, et al. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer Cell*. 2018;33(5):853-861.e4. doi:10.1016/j.ccell.2018.04.001

28. R: a language and environment for statistical computing [computer program]. Vienna, Austria: the R Foundation for Statistical Computing; 2011.

29. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31. doi:10.1186/1471-2288-6-31

30. Kefford R, Ribas A, Hamid O, et al. Clinical efficacy and correlation with tumor PD-L1 expression in patients (pts) with melanoma (MEL) treated with the anti-PD-1 monoclonal antibody MK-3475. *J Clin Oncol*. 2014;32(15)(suppl):3005-3005. doi:10.1200/jco.2014.32.15_suppl.3005

31. Puzanov I, Dummer R, Schachter J, et al. Efficacy based on tumor PD-L1 expression in KEYNOTE-002, a randomized comparison of pembrolizumab (PEMBRO; MK-3475) versus chemotherapy in patients (pts) with ipilimumab-refractory (IPI-R) advanced melanoma (MEL). *J Clin Oncol*. 2015;33(15)(suppl):3012-3012. doi:10.1200/jco.2015.33.15_suppl.3012

32. 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(4):383-390. doi:10.1200/JCO.2016.71.8023

33. Wolchok JD, Chiarion-Sileni V, Gonzalez R, et al. Overall survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med*.

2017;377(14):1345-1356. doi:10.1056/NEJMoa1709684

34. Robert C, Long GV, Brady B, et al. Nivolumab in previously untreated melanoma without *BRAF* mutation. *N Engl J Med*. 2015;372(4):320-330. doi:10.1056/NEJMoa1412082

35. 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(10027):1540-1550. doi:10.1016/S0140-6736(15)01281-7

36. Horn L, Spigel DR, Vokes EE, et al. Nivolumab versus docetaxel in previously treated patients with advanced non-small-cell lung cancer: two-year outcomes from two randomized, open-label, phase III trials (CheckMate 017 and CheckMate 057). *J Clin Oncol*. 2017;35(35):3924-3933. doi:10.1200/JCO.2017.74.3062

37. Rittmeyer A, Barlesi F, Waterkamp D, et al; OAK Study Group. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *Lancet*. 2017;389 (10066):255-265. doi:10.1016/S0140-6736(16)32517-X

38. Balar AV, Castellano D, O'Donnell PH, et al. First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol*. 2017;18(11):1483-1492. doi:10.1016/S1470-2045(17)30616-2

39. Bellmunt J, de Wit R, Vaughn DJ, et al; KEYNOTE-045 Investigators. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N Engl J Med*. 2017;376(11):1015-1026. doi:10.1056/NEJMoa1613683

40. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *Lancet Oncol*. 2017;18 (3):312-322. doi:10.1016/S1470-2045(17)30065-7

41. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet*. 2016;387(10031):1909-1920. doi:10.1016/S0140-6736(16)00561-4

42. Balar AV, Galsky MD, Rosenberg JE, et al; IMvigor210 Study Group. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet*. 2017;389(10064):67-76. doi:10.1016/S0140-6736(16)32455-2

43. Powles T, O'Donnell PH, Massard C, et al. Efficacy and safety of durvalumab in locally advanced or metastatic urothelial carcinoma: updated results from a phase 1/2 open-label study. *JAMA Oncol*. 2017;3(9):e172411. doi:10.1001/jamaoncol.2017.2411

44. Patel MR, Ellerton J, Infante JR, et al. Avelumab in metastatic urothelial carcinoma after platinum failure (JAVELIN Solid Tumor): pooled results from two expansion cohorts of an open-label, phase 1 trial. *Lancet Oncol*. 2018;19(1):51-64. doi:10.1016/S1470-2045(17)30900-2

45. Chow LQM, Haddad R, Gupta S, et al. Antitumor activity of pembrolizumab in biomarker-unselected patients with recurrent and/or metastatic head and neck squamous cell carcinoma: results from the phase Ib KEYNOTE-012 expansion cohort. *J Clin Oncol*. 2016;34(32):3838-3845. doi:10.1200/JCO.2016.68.1478

46. 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(19):1856-1867. doi:10.1056/NEJMoa1602252

47. Fuchs CS, Doi T, Jang RW, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. *JAMA Oncol*. 2018;4(5):e180013. doi:10.1001/jamaoncol.2018.0013

48. Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet Oncol*. 2017;18(9):1182-1191. doi:10.1016/S1470-2045(17)30422-9

49. El-Khoueiry AB, Sangro B, Yau T, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet*. 2017;389 (10088):2492-2502. doi:10.1016/S0140-6736(17)31046-2

50. Kaufman HL, Russell J, Hamid O, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. *Lancet Oncol*. 2016;17(10):1374-1385. doi:10.1016/S1470-2045(16)30364-3

51. Chung HC, Schellens JHM, Delord J-P, et al. Pembrolizumab treatment of advanced cervical cancer: updated results from the phase 2 KEYNOTE-158 study. *J Clin Oncol*. 2018;36(15) (suppl):5522-5522. doi:10.1200/JCO.2018.36.15_suppl.5522

52. 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(7):633-641. doi:10.1200/JCO.2017.75.3384

53. Galsky M, Saci A, Szabo P, et al. Impact of tumor mutation burden on nivolumab efficacy in second-line urothelial carcinoma patients: exploratory analysis of the phase II CheckMate 275 study. *Ann Oncol*. 2017;28(suppl 5):v295-v329. doi:10.1093/annonc/mdx371.003

54. Hugo W, Zaretsky JM, Sun L, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell*. 2016;165(1):35-44. doi:10.1016/j.cell.2016.02.065

55. Riaz N, Havel JJ, Makarov V, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell*. 2017;171(4):934-949.e16. doi:10.1016/j.cell.2017.09.028

56. Johnson DB, Frampton GM, Rioth MJ, et al. Targeted next generation sequencing identifies markers of response to PD-1 blockade. *Cancer Immunol Res*. 2016;4(11):959-967.

57. Ribas A, Robert C, Hodi FS, et al. Association of response to programmed death receptor 1 (PD-1) blockade with pembrolizumab (MK-3475) with an interferon-inflammatory immune gene signature. *J Clin Oncol*. 2015;33(15)(suppl):3001-3001. doi:10.1200/jco.2015.33.15_suppl.3001

58. Karachaliou N, Crespo G, Aldeguer E, et al. Interferon-gamma (INFG), an important marker of response to immune checkpoint blockade (ICB) in non-small cell lung cancer (NSCLC) and melanoma patients. *J Clin Oncol*. 2017;35(15)(suppl):11504. doi:10.1200/JCO.2017.35.15_suppl.11504

59. Wang L, Saci A, Szabo PM, et al. EMT- and stroma-related gene expression and resistance to PD-1 blockade in urothelial cancer. *Nat Commun*. 2018;9(1):3503. doi:10.1038/s41467-018-05992-x

60. Althammer S, Tan TH, Spitzmüller A, et al. Automated image analysis of NSCLC biopsies to predict response to anti-PD-L1 therapy. *J Immunother Cancer*. 2019;7(1):121. doi:10.1186/s40425-019-0589-x

61. Mazzaschi G, Madeddu D, Falco A, et al. Low PD-1 expression in cytotoxic CD8+ tumor-infiltrating lymphocytes confers an immune-privileged tissue microenvironment in NSCLC with a prognostic and predictive value. *Clin Cancer Res*. 2018;24(2):407-419. doi:10.1158/1078-0432.CCR-17-2156

62. Danilova L, Wang H, Sunshine J, et al. Association of PD-1/PD-L axis expression with cytolytic activity, mutational load, and prognosis in melanoma and other solid tumors. *Proc Natl Acad Sci U S A*. 2016;113(48):E7769-E7777. doi:10.1073/pnas.1607836113

63. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;372(26):2509-2520. doi:10.1056/NEJMoa1500596

64. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248-262. doi:10.1016/j.celrep.2016.12.019

65. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098

66. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Publication bias. In: Borenstein M, Hedges LV, Higgins JPT, Rothstein HR, eds. *Introduction to Meta-Analysis*. Chichester, West Sussex, United Kingdom: John Wiley & Sons; 2009. doi:10.1002/9780470743386.ch30

67. Frati A, Chereau E, Coutant C, et al. Comparison of two nomograms to predict pathologic complete responses to neoadjuvant chemotherapy for breast cancer: evidence that HER2-positive tumors need specific predictors. *Breast Cancer Res Treat*. 2012;132(2):601-607. doi:10.1007/s10549-011-1897-0

68. English PA, Williams JA, Martini JF, Motzer RJ, Valota O, Buller RE. A case for the use of receiver operating characteristic analysis of potential clinical efficacy biomarkers in advanced renal cell carcinoma. *Future Oncol*. 2016;12(2):175-182. doi:10.2217/fon.15.290

69. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203-211.

70. Berry S, Giraldo N, Nguyen P, et al. 33rd annual meeting & pre-conference programs of the Society for Immunotherapy of Cancer (SITC 2018): P128 multiplexed immunofluorescent assay development for study of the PD-1/PD-L1 checkpoint in the tumor immune microenvironment (TIME). *J Immunother Cancer*. 2018;6(1):114.

71. Feng Z, Bethmann D, Kappler M, et al. Multiparametric immune profiling in HPV-oral squamous cell cancer. *JCI Insight*. 2017;2(14):93652. doi:10.1172/jci.insight.93652

72. Tsujikawa T, Kumar S, Borkar RN, et al. Quantitative multiplex immunohistochemistry reveals myeloid-inflamed tumor-immune complexity associated with poor prognosis. *Cell Rep*. 2017;19(1):203-217. doi:10.1016/j.celrep.2017.03.037

73. Remark R, Merghoub T, Grabe N, et al. In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide. *Sci Immunol*. 2016;1(1):aaf6925. doi:10.1126/sciimmunol.aaf6925

74. Giesen C, Wang HA, Schapiro D, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods*. 2014;11(4):417-422. doi:10.1038/nmeth.2869

75. Schulz D, Zanotelli VRT, Fischer JR, et al. Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry. *Cell Syst*. 2018;6(4):531. doi:10.1016/j.cels.2018.04.004

76. Keren L, Bosse M, Marquez D, et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell*. 2018;174(6):1373-1387.e19. doi:10.1016/j.cell.2018.08.039

77. Angelo M, Bendall SC, Finck R, et al. Multiplexed ion beam imaging of human breast tumors. *Nat Med*. 2014;20(4):436-442. doi:10.1038/nm.3488

78. Velcheti V, Schalper KA, Carvajal DE, et al. Programmed death ligand-1 expression in non-small cell lung cancer. *Lab Invest*. 2014;94(1):107-116. doi:10.1038/labinvest.2013.130

79. Spranger S, Luke JJ, Bao R, et al. Density of immunogenic antigens does not explain the presence or absence of the T-cell-inflamed tumor microenvironment in melanoma. *Proc Natl Acad Sci U S A*. 2016;113(48):E7759-E7768. doi:10.1073/pnas.1609376113

80. Jain RK, Lee JJ, Ng C, et al. Change in tumor size by RECIST correlates linearly with overall survival in phase I oncology studies. *J Clin Oncol*. 2012;30(21):2684-2690. doi:10.1200/JCO.2011.36.4752