The Immunohistochemical Expression of Programmed Death Ligand 1 (PD-L1) Is Affected by Sample Overfixation

Angels Barberà, MNurs,† Ruth Marginet Flinch, APT,† Montserrat Martin, APT,† Jose L. Mate, PhD,† Albert Oriol, MD,‡ Fina Martínez-Soler, PhD,* Tomas Santalucia, PhD,* and Pedro L. Fernández, PhD†*

Abstract: Humanized antibodies targeting programmed death receptor 1 (PD-1) or its ligand (PD-L1) have been approved for the treatment of different cancers. Some of these antibodies show a correlation between the tissue expression of PD-L1 and response. Evaluation of PD-L1 expression presents multiple challenges, but some preanalytical issues such as tissue fixation have been scarcely evaluated. With the hypothesis that immunohistochemical staining of PD-L1 may be impacted by the time of specimen fixation, we evaluated differences in its expression in tonsil samples exposed to predefined fixation times. Random nontumoral tonsillectomy specimens were blindly evaluated in tissue microarray slides after staining with SP142 and SP263 antibodies. With fixation times ranging from 12 to 72 hours, between 2.8% and 6.1% of the samples were considered to be suboptimally stained, with no differences between the 2 antibodies within these fixation times. A significantly higher proportion of samples exposed to a fixation time of 96 hours presented suboptimal immunostaining (15.6%, $P < 0.0001$). In addition, suboptimally stained spots were 20.8% using SP142 and 10.4% using SP263 after 96 hours of fixation ($P = 0.046$). In conclusion, the quality of staining for PD-L1 in tonsil samples decreased with overfixation of the specimen at times > 72 hours. Samples exposed to formaldehyde for longer periods presented suboptimal results for both clones, but the SP142 antibody presented a significantly lower tolerance to formalin overexposure than SP263. These results indicate the relevance of a controlled preanalytical processing of samples and particularly the length of fixation of tumor specimens.

Key Words: PD-L1 immunochemistry, preanalytical processes, fixation times

(*Appl Immunohistochem Mol Morphol* 2020;00:000–000)

Received for publication February 3, 2020; accepted February 11, 2020.
From the *Fundamental Care and Medical-Surgical Nursing Department, School of Nursing, Faculty of Medicine and Health Sciences, Barcelona University; †Pathology Department, Germans Trias i Pujol Hospital and Institute, Faculty of Medicine and Health Sciences, Universitat Autònoma de Barcelona; and ‡Josep Carreras Leukemia Research Institute, Barcelona, Spain.
The authors declare no conflict of interest.
Reprints: Angels Barberà, MNurs, Hospital Germans Trias i Pujol, Carretera del Canyet, Badalona, Barcelona 08916, Spain (e-mail: mabarbera.germanstrias@gencat.cat).

The concept of immune surveillance and editing of tumors is now well-accepted, particularly after the emergence of immunotherapy as an effective strategy to treat several types of solid tumors with very good results.[1] Programmed death receptor 1 (PD-1) is normally expressed on the surface of T and B cells, natural killer cells, dendritic cells, and macrophages and plays a key role in the physiological restriction of immune-mediated tissue damage secondary to inflammation.[2] Therapeutic blockade of PD-1/programmed death ligand 1 (PD-L1) interaction can restore the function of exhausted T cells. In recent years, several humanized antibodies targeting PD-1 or PD-L1 have been approved to treat different cancers, including non–small cell lung[3] and urothelial carcinomas,[4] melanoma,[5] and other solid and hematological malignancies.[6–8] For some of these tumors, there seems to be a correlation between tissue expression of PD-L1 and response rate in patients.[9] Different PD-L1 immunohistochemistry kits, clones, and platforms have been validated in recent years. Nevertheless, its evaluation is still challenging. PD-L1 expression is highly heterogenous not only among different tumor types but also among malignant and inflammatory cells within a tumor. Likewise, the expression of the receptor molecule is dynamic and extremely variable among immune effector cells within a same neoplastic tumor. From the point of view of a pathologist, standardization of diagnostic criteria requires establishing cutpoints for each antibody and platform.[10] In addition to the inherent tumor heterogeneity and the variety of methods to evaluate samples, several preanalytical and analytical procedures may have an impact on the quality of the final specimens. While most of the analytical challenges in the evaluation of PD-L1 expression have been extensively analyzed,[11] preanalytical issues such as optimal tissue fixation have been scarcely evaluated. The primary objective of this study was to evaluate differences in the assessment of PD-L1 expression in tonsil samples exposed to predefined fixation times. In addition, we compared the robustness and reproducibility of 2 commonly used antibodies along the different fixation times.

METHODS

Random consecutive nontumoral tonsillectomy samples were used for the study. The samples were processed fresh. Each sample was coded, and up to 10 sections were obtained from each sample whenever possible. Sections had a 1 cm^2 area and were 0.5 cm thick. Twenty-two samples were used to obtain at least 16 evaluable sections for each fixation time.

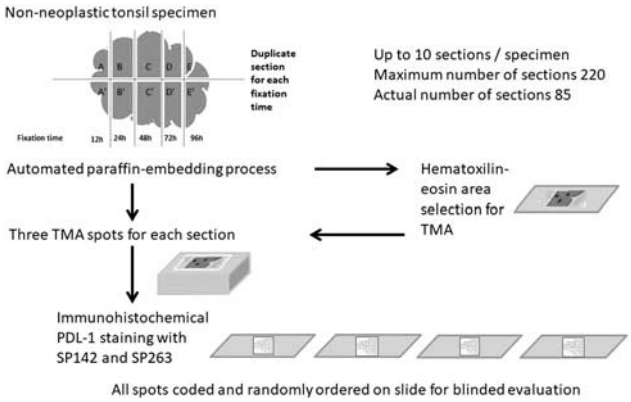

FIGURE 1. Specimen processing, allocation, and coding of sections. PD-L1 indicates programmed death ligand 1;TMA, tissue microarray.

At least 2 sections for each fixation time were obtained of each sample to evaluate intrasample variability in the whole sample set (Fig. 1). Duplicate sections were formalin-fixed at each of the selected fixation times. Fixation times were randomly skipped for samples in which <10 sections were obtainable. All samples were fixed in neutral buffered formalin pH 7.2 to 7.4 (Diapath F0048) with a formalin/sample volume proportion of 10:1. Predefined fixation times of 12, 24, 48, 72, and 96 hours (including the final 120 minutes of in-processor fixation) were applied. Following fixation, fragments underwent standard clinical processing in paraffin blocks. The dehydration and paraffin-embedding process was automated using Leica ASP3000S equipment and following standard routine laboratory procedures. An experienced pathologist assessed cellularity and histologic preservation in the hematoxylin and eosin–stained tumor samples to identify representative areas to be selected for tissue microarrays (TMAs). Three 3 mm punches of the selected areas were obtained for each representative block. The tissue cores were inserted into 4 TMA paraffin 6×5 recipient blocks using a Beecham arrayer. Each tissue core was assigned a unique tissue microarray location number that was linked to an external database.

Immunohistochemistry

TMA slides were processed in a BenchMark ULTRA platform instrument (Ventana Medical Systems, Roche, Tucson, AZ) and stained with SP142 and SP263 antibodies (Ventana, assay refs 740-4859 and 790-4905), both of which were prediluted following the manufacturer's instructions. Interpretation of immunohistochemical staining was performed independently by 2 experienced pathologists (P.L.F. and J.L.M.). The 2 observers evaluated the expression of immunohistochemical markers in epithelial and immune cells by light microscopy. Spots in the TMA were allocated randomly to ensure pathologists were blind to any information, including tonsil origin, antibody, and fixation time, when assessing each spot. Readers were asked to evaluate each spot as "optimally stained" or "suboptimally stained" according to objective prespecified criteria.[12] An "optimally stained" specimen was defined as including moderate to intense staining (+1 or +2) of macrophagic cells and T lymphocytes in germinal centers, intense staining (+2) of some epithelial squamous cells and negativity in cells of interfollicular regions and superficial epithelium, and no background staining. Samples needed to meet both positive and negative criteria in at least 90% of the observed dot cellularity to be considered as optimally stained, and, consequently, the main variable to be analyzed was dichotomized (optimal/suboptimal). A separate training set was observed and revised before evaluation of the final study set. The training set was used to ensure <10% intraobserver and interobserver discrepancies before the core set was evaluated. Training set results were reviewed and discussed for case that showed discrepancies among the 2 pathologists or within evaluations of the same pathologist. After discrepancies were discussed to achieve maximal consensus among observers with regard to borderline cases, and when <10% intrareader and interreader discrepancies were achieved in the training set, the final sample set was released for evaluation.

Statistical Analysis

Interreader, intrareader, intersample, and intrasample concordance rates were described and analyzed using the κ-Cohen method in the final sample set.[13] Each reading was classified as optimal if meeting previously defined criteria or suboptimal if criteria were not met for any reason. Interreader discrepancies refer to a same spot (same sample, antibody, and fixation time) evaluated differently by the 2 readers. An intrareader discrepancy was considered when 1 of the 3 spots from a given sample (same block, antibody, and fixation time) was evaluated differently from the other 2 by the same reader. Intrasample discrepancies refer to differences between 2 sections of a same tonsil (same antibody, observer, and fixation time). Finally, samples from different tonsils were matched (according to observer, antibody, and fixation time) and compared for discrepancies. Intersample comparisons were used to ensure acceptable uniformity among different tonsils if the κ value was similar to that of intrasample comparisons. Tests on the concordance rate among PD-L1 antibodies, among fixation times, and among fixation times stratified by antibody were performed using the χ^2 statistic. Differences were considered to be statistically significant when P-value <0.05. The R-package, version 3.0.1. was used for all statistical analyses (R Foundation for Statistical Computing, Vienna, Austria).[14]

Ethical Aspects

This study was approved by the Institutional Review Board and Ethics Committee. Patients were asked to voluntarily donate tonsil samples before surgery and signed an institutionally approved informed consent form. All the samples used in the study were irreversibly dissociated from tonsillectomized patients' data.

RESULTS

Twenty-two non-neoplastic tonsillectomy specimens were processed, and a total of 92 evaluable pieces were obtained (median: 4 pieces/specimen, range: 2 to 10), yielding 276 representative TMA spots that made up the final sample. A training set of 660 readings was used to determine final intrasample, intersample, intraobserver, and interobserver

TABLE 1. Reproducibility of Readings for Related Spots

	n (%)			
	Positive	Negative	Discrepancy	κ Value
Intrasample	551	52 (8)	57 (9)	0.61
Intersample*	93	17 (14)	11 (9)	0.70
Intraobserver†	304	24 (7)	26 (7)	0.61
Interobserver	551	37 (6)	33 (5)	0.66

Not all tonsils provided enough material for all fixation times. Missing values were excluded pairwise. The κ value for intrasample reproducibility refers to results of spots of matched sections of the same tonsil (same observer, antibody, and fixation time). Interobserver reproducibility refers to same spot, as evaluated by each observer. Intrasample reproducibility spots were matched according to same observer, antibody, and fixation time but different sample of origin (*only spots that could be matched were compared). Intraobserver reproducibility was evaluated taking into account the 3 spots of a same block (same antibody and fixation time).

†The triplicate reading was considered discrepant if one spot was discrepant; for this reason, the κ value was expected to be lower than the κ value for interobserver readings.

reproducibility. The results are shown in Table 1. Less than 10% of discrepancies were observed. All of the observations evaluated as suboptimally stained were related to false negativity in cells expected to be positive. No sample was considered to be suboptimally stained because of positivity in cells expected to be negative or because of background staining. κ values were within the range of 0.6 for all comparisons, including intraobserver variability with which the κ value was based on a triplicate observation.

When considering the 2 PD-L1 monoclonal antibodies, SP142 and SP263, 45 of 552 observations (8.15%) were rated as suboptimally stained compared with 26 of 552 (4.71%), respectively, with the difference being statistically significant ($P = 0.021$). However, the concordance among monoclonal PD-L1 antibodies was not uniform across the different fixation times. The proportion of suboptimally stained samples ranged between 2.77% and 6.14% for samples fixed for 12 to 72 hours, and no statistically significant differences were found between antibodies for any of the fixation time periods (Fig. 2). Samples exposed to a fixation time of 96 hours presented suboptimal immunostaining in 30 of 192 cases (15.63%). The proportion of suboptimally stained samples with this fixation time was significantly higher than for all the other time periods ($P < 0.0001$). In addition, statistically significant differences between antibodies were obtained among samples in the 96-hour fixation group, with 20.83% of suboptimally stained spots using SP142 and 10.42% using SP263 ($P = 0.046$). The quality of the staining was comparable for fixation times of 12 to 72 hours. There were no statistically significant differences between SP142 and SP263 for any individual fixation time <96 hours. Figure 3 shows the impact of fixation time on the quality of immunostaining for tonsil #8. In addition, intrareader and interreader concordance was lower in suboptimally stained spots, with agreement between readers being 60.5%.

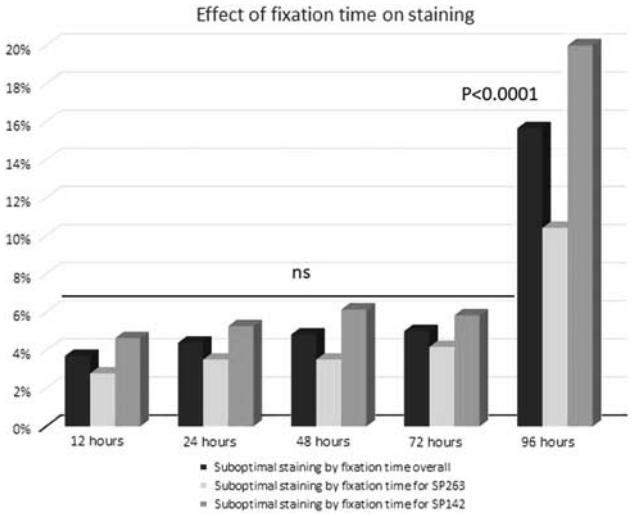

FIGURE 2. Proportion of programmed death ligand 1 readings evaluated as suboptimally stained according to antibody and fixation time. ns indicates non significant.

DISCUSSION

In this study, we hypothesized that the reproducibility and accuracy of the interpretation of immunohistochemistry staining for PD-L1 may be affected by the sample fixation time.

The interaction of PD-1 and PD-L1 has proved to be a major mechanism for cancer cells to avoid antitumor immune responses.[1,2] A variable expression of PD-L1 and PD-1 has been described in lung[3] and urinary bladder cancer[4] and malignant melanoma[5] among a growing number of other neoplastic cells. Targeted therapy for PD-L1-expressing neoplastic conditions has represented a major step forward in the treatment of several malignancies in recent years,[15] and knowledge of the expression of this marker by cancer cells has become a central diagnostic procedure in tailored therapy planning. Although the expression of PD-L1 in formalin-fixed and paraffin-embedded specimens can be easily evaluated using immunohistochemistry, the prognostic and therapeutic implications of this evaluation are related to its precision, reliability, and, ultimately, standardization.[16] Adequate diagnostic evaluation of PD-L1 expression is challenged by several tumor-related issues. There is an important biological variability among patients: expression rates in lung cancer vary from 24% to 60%,[17] and there may also be extreme heterogeneity within a tumor and even within the same tumor sample.[18] In addition, PD-L1 expression may be dynamic, as it is inducible by inflammatory cytokines released into the microenvironment, and it may even be upregulated or downregulated by different antineoplastic agents.[19] This is particularly important when considering its expression to define treatment at the time of relapse. Furthermore, PD-L1 is expressed by T cells, including Tregs and natural killer cells, and some clinical studies have suggested that this expression in tumor-infiltrating immune cells may also be associated with response to specific treatment.[20] Some studies have specifically revealed low concordance in reporting immune cell scores.[21] Beyond the biological sample itself, several analytical aspects have been a matter of concern. The progressive adoption of standardized procedures and consensus assessment criteria has been difficult, as different PD-L1

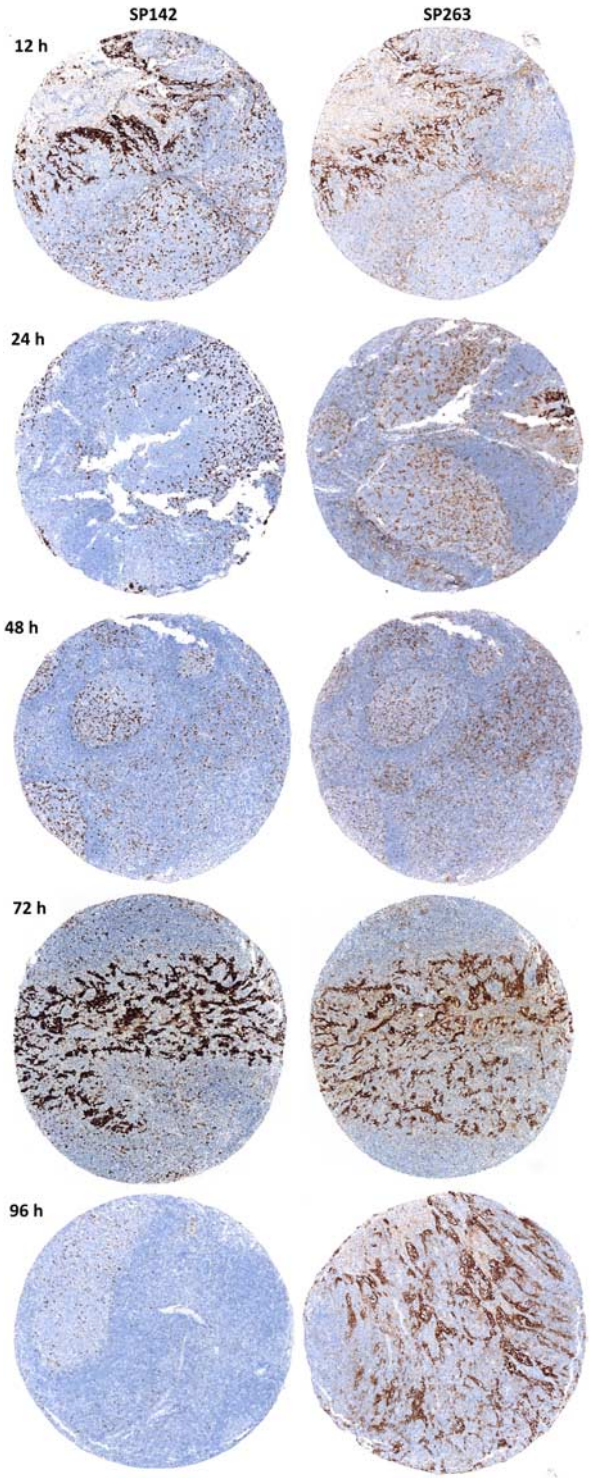

FIGURE 3. Range of programmed death ligand 1 expression in a tonsil specimen for each antibody and fixation time.

antibodies coupled with partner immunohistochemistry platforms are routinely used in clinical practice,[22,23] and the comparability of the various PD-L1 assays and platforms is far from perfect.[24] Different clones of monoclonal antibodies raised against the same protein are, in fact, specific for different protein epitopes, and immunohistochemical interpretation may be complicated by this fact. Visualization of the bound primary antibody on the tissue section is dependent on the detection system used, and one test may not necessarily perform identically to others. In addition, individual laboratories may choose to use 1 or more different antibodies in a single platform to optimize logistics or to meet the preferences of their requesting oncologists. Even in the most standardized setting, the chance of obtaining false-negative results varies greatly according to the cutoff values used to classify a sample as negative or positive.[25] The coexistence of different clones, staining platforms, and scoring criteria, have been a matter of concern and have been extensively addressed in recent years.[26]

Less consideration has been given to the potential effect of preanalytical issues (eg, epitope stability in tissue blocks and stored sections) on PD-L1 expression. Preanalytical issues may have a particularly dramatic impact when small diagnostic samples are tested. In the present study, we focused on a scarcely explored area that may add to the current uncertainties and complexities of PD-L1 assessment. Formalin-fixation is a critical part of the preanalytical processing of tissue samples that may affect immunohistochemical results, and, consequently, it may also affect the reliability of PD-L1 assessment. Thus, our study provides support to define a maximum range of fixation time to consider a tumor sample as evaluable for PD-L1. It is likely that this predefined range should be even stricter when assessing diagnostic tumor samples, with which the heterogeneity of expression is expected to be much higher than tonsil specimens, and certain small size samples (ie, bronchoscopic biopsy specimens) may be even more sensitive to excessive fixation.

Previous studies have shown that evaluation of immunohistochemically stained tumoral tissue sections are also characterized by significant intraobserver and interobserver variability.[27] Tonsil tissue is the standard control tissue for PD-L1 evaluation, its interpretation is well defined, and it is assumed to be minimally affected by variability. Our evaluation of intrasample and intersample correlations confirmed that tonsil tissue has a reliable reproducibility when processed under identical conditions. However, intraobserver and interobserver concordance significantly improved when clear-cut evaluation criteria were used to assess TMA spots, and only after training was concordance of 90% achieved between experienced observers.

Several study limitations are particularly relevant. Although the dichotomic endpoint (optimal vs. suboptimal) was strictly defined, a number of borderline cases were noted by the observers, both in the training and in the study set. Spots with scarce cellularity and evaluation of intermediate positive results were particularly challenging. In fact, evaluations showing a lack of intraobserver and interobserver concordance occurred in these situations. The categorization of fixation times into predefined levels and dichotomized "optimal" and "suboptimal" evaluations should not obscure the fact that the relationship of overfixation and the rate of false-negative samples should be considered as a continuous function.

A second limitation is that results in tonsil tissue are not readily translatable to the tumor specimens that are of real relevance in the clinical setting. PD-L1 expression in tonsil tissue is uniform both in extension and intensity throughout the tissue. Indeed, the existence of both intense positive cells and frequent negative elements allows clear-cut evaluation of the samples and a relatively straightforward assessment of the impact of fixation. PD-L1 expression in lung and other neoplastic tissues is much more heterogenous, and, in fact, relevant discordance rates between TMA samples and whole tissue sections have been reported.[28] Moreover, the small size of tumor biopsy samples such as those of the lung obtained through bronchoscopic procedures can preclude optimal PD-L1 evaluations in some instances. Here, we have provided clear evidence that excessively prolonged fixation times may produce a relevant number of suboptimal results. Therefore, our results strongly recommend to follow strictly the Food and Drug Administration (FDA)-cleared PD-L1 (SP142 and SP263) manufacturer recommendations with regard to fixation times for tonsil and placenta specimens. However, extrapolation of such results to specific tumor tissues should be made with caution. Acceptable fixation time ranges for different tumor types and specimen characteristics are likely to be variable taking into consideration the higher heterogeneity in PD-L1 expression and may require the specific evaluation of different tumoral tissues or even specimen volumes. It is much likely that smaller and more heterogenous tissue types require an even stricter control of fixation times, and we are currently performing a study in this respect. Finally, several monoclonal antibodies are currently available for the evaluation of PD-L1 expression. A full comparison of the performance of the increasing number of antibodies and platforms was beyond the scope of the present study. Other authors have already addressed this issue[26,29]; however, our immunohistochemical analysis was carried out under very controlled preanalytical conditions with 2 different antibodies that are routinely used in our laboratory, and we proved that, while they had a similar performance at optimal fixation times, they were not equally robust when the samples were exposed to overfixation. Consequently, studies that seek to evaluate the performance of different antibodies should also take into consideration that, despite similar performances under optimal conditions, they may have significantly different reliability when samples are subjected to less adequate processing.

In conclusion, the quality of staining for PD-L1 in tonsil samples varies with time of specimen fixation. High-quality staining was obtained for fixation times up to 72 hours, but a significant proportion of samples exposed to formaldehyde for a longer period presented suboptimal results for both of the clones tested. The SP142 antibody presented a significantly lower tolerance to formalin overexposure than SP263. These results indicate the relevance of controlled preanalytical processing of samples, particularly with regard to time of fixation of tumor specimens, with a potential impact on the prognostic and therapeutic approaches derived from the evaluation of this biomarker.

ACKNOWLEDGMENTS

The authors thank the technical and auxiliary staff in the Pathology Department who managed the collection and processing of samples and the institutional support provided by the Nursing Administration of the Hospital Germans Trias i Pujol.

REFERENCES

1. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science*. 2011;331:1565–1570.
2. Torphy RJ, Schulick RD, Zhu Y. Newly emerging immune checkpoints: promises for future cancer therapy. *Int J Mol Sci*. 2017;18:2642.
3. Gandhi L, Rodriguez-Abreu D, Gadgeel S, et al. Pembrolizumab-plus chemotherapy in metastatic non-small-cell lung cancer. *N Engl J Med*. 2018;378:2078–2092.
4. Sharma P, RetzM, Siefker-Radtke A, et al. Nivolumab inmetastaticurothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *Lancet Oncol*. 2017;18:312–322.
5. Postow MA, Chesney J, Pavlick AC, et al. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med*. 2015;372:2006–2017.
6. Seiwert TY, Burtness B, Mehra R, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinomaof the head and neck (KEYNOTE-012): an open-label, multicentre, phase1b trial. *Lancet Oncol*. 2016;17:956–965.
7. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade withnivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med*. 2015;372:311–319.
8. Cimino-Mathews A, Thompson E, Taube JM, et al. PD-L1 (B7-H1) expression and the immune tumor microenvironment in primary and metastatic breast carcinomas. *Hum Pathol*. 2016;47:52–63.
9. Pan ZK, Ye F, Wu X, et al. Clinicopathological and prognostic significance of programmed cell death ligand1 (PD-L1) expression in patients with non-small cell lung cancer: a meta-analysis. *J Thorac Dis*. 2015;7:462–470.
10. Vennapusa B, Baker B, Kowanetz M, et al. Development of a PD-L1 complementary diagnostic immunohistochemistry assay (SP142) for atezolizumab. *Appl Immunohistochem Mol Morphol*. 2019;27:92–100.
11. Parra ER, Villalobos P, Mino B, et al. Comparison of different antibody clones for immunohistochemistry detection of programmed cell death ligand 1 (PD-L1) on non-small cell lung carcinoma. *Appl Immunohistochem Mol Morphol*. 2018;26:83–93.
12. Ventana Medical Systems Inc. Ventana PD-L1 (SP142) assay. Interpretation guide P/N 1014987EN; 2016. Available at: www.ventana.com. Accessed January 20, 2020.
13. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257–268.
14. R Core Team. *R: A Language and Environment for Statistical Computing 2019*. Vienna, Austria: R Foundation for Statistical Computing.
15. Wei SC, Duffy CR, Allison JP. Fundamental mechanisms of immune checkpoint blockade therapy. *Cancer Discov*. 2018;8:1069–1086.
16. Hersom M, Jørgensen JT. Companion and complementary diagnostics-focus on PD-L1 expression assays for PD-1/PD-L1 checkpoint inhibitors in non-small cell lung cancer. *Ther Drug Monit*. 2018;40:9–16.
17. Song Z, Yu X, Cheng G, et al. Programmed death-ligand 1 expression associated with molecular characteristics in surgically resected lung adenocarcinoma. *J Transl Med*. 2016;14:188.
18. Nakamura S, Hayashi K, Imaoka Y, et al. Intratumoral heterogeneity of programmed cell death ligand-1 expression is common in lung cancer. *PLoS One*. 2017;12:e0186192.
19. Champain S, Ileana E, Giaccone G, et al. Incorporating immune-checkpoint inhibitors into systemic therapy of NSCLC. *J Thorac Oncol*. 2014;9:144–153.

20. Kim HR, Ha SJ, Hong MH, et al. PD-L1 expression on immune cells, but not on tumor cells, is a favorable prognostic factor for head and neck cancer patients. *Sci Rep*. 2016;6:36956.
21. Rimm DL, Han G, Taube JM, et al. A prospective, multiinstitutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-1 expression in non–small cell lung cancer. *JAMA Oncol*. 2017;3:1051–1058.
22. Koppel C, Schwellenbach H, Zielinski D, et al. Optimization and validation of PD-L1 immunohistochemistry staining protocols using the antibody clone 28-8 on different staining platforms. *Mod Pathol*. 2018;31:1630–1644.
23. Adam J, Le Stang N, Rouquette I, et al. Multicenter harmonization study for PD-L1 IHC testing in non-small-cell lung cancer. *Ann Oncol*. 2018;29:953–958.
24. Chakravarti N, Prieto VG. Predictive factors of activity of anti-programmed death-1/programmed death ligand-1 drugs: immunohistochemistry analysis. *Transl Lung Cancer Res*. 2015;4:743–751.
25. Ratcliffe MJ, Sharpe A, Midha A, et al. Agreement between programmed cell death ligand-1 diagnostic assays across multiple protein expression cutoffs in non-small cell lung cancer. *Clin Cancer Res*. 2017;23:3585–3591.
26. Büttner R, Gosney JR, Skov BG, et al. Programmed death-ligand 1 immuno-histochemistry testing: a review of analytical assays and clinical implementation in non-small-cell lung cancer. *J Clin Oncol*. 2017;35:3867–3876.
27. Chang S, Park HK, Choi YL, et al. Interobserver reproducibility of pd-L1 biomarker in non-small cell lung cancer: a multi-institutional study by 27 pathologists. *J Pathol Transl Med*. 2019;53:347–353.
28. Li C, Huang C, Mok TS, et al. Comparison of 22C3 PD-L1 expression between surgically resected specimens and paired tissue microarrays in non-small cell lung cancer. *J Thorac Oncol*. 2017;12:1536–1543.
29. Gaule P, Smithy JW, Toki M, et al. A quantitative comparison of antibodies to programmed cell death 1 ligand 1. *JAMA Oncol*. 2017;3:256–259.