ORIGINAL ARTICLE

PD-L1 Expression Heterogeneity in Non–Small Cell Lung Cancer: Defining Criteria for Harmonization between Biopsy Specimens and Whole Sections

Enrico Munari, MD, PhD,[a,b,*] Giuseppe Zamboni, MD,[a,b] Gianluigi Lunardi, MPharm,[c] Luigi Marchionni, MD, PhD,[d] Marcella Marconi, BS,[a] Marco Sommaggio, BS,[a] Matteo Brunelli, MD, PhD,[b] Guido Martignoni, MD,[b,e] George J. Netto, MD,[f] Mohammad O. Hoque, PhD,[g] Francesca Moretta, MD,[h] Maria Cristina Mingari, PhD,[i] Maria Cristina Pegoraro, MD,[j] Alessandro Inno, MD,[c] Simona Paiano, MD,[k] Alberto Terzi, MD,[l] Alberto Cavazza, MD,[m] Giulio Rossi, MD,[n] Francesca Romana Mariotti, PhD,[o] Paola Vacca, PhD,[o] Lorenzo Moretta, MD,[o] Giuseppe Bogina, MD[a]

[a]*Department of Pathology, Sacro Cuore Don Calabria Hospital, Negrar, Verona, Italy*
[b]*Department of Diagnostics and Public Health, University of Verona, Verona, Italy*
[c]*Department of Oncology, Sacro Cuore Don Calabria Hospital, Negrar, Verona, Italy*
[d]*Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland*
[e]*Department of Pathology, Pederzoli Hospital, Peschiera del Garda, Verona, Italy*
[f]*Department of Pathology, The University of Alabama at Birmingham, Birmingham, Alabama*
[g]*Department of Otolaringology, Johns Hopkins University School of Medicine, Baltimore, Maryland*
[h]*Department of Laboratory Medicine, Sacro Cuore Don Calabria Hospital, Negrar, Verona, Italy*
[i]*Department of Experimental Medicine, University of Genova, Genoa, Italy*
[j]*Department of Oncology, Pederzoli Hospital, Peschiera del Garda, Verona, Italy*
[k]*Department of Pulmonology, Sacro Cuore Don Calabria Hospital, Negrar, Verona, Italy*
[l]*Department of Thoracic Surgery, Sacro Cuore Don Calabria Hospital, Negrar, Verona, Italy*
[m]*Department of Pathology, Arcispedale S. Maria Nuova/IRCCS, Reggio Emilia, Italy*
[n]*Department of Pathology, AUSL della Romagna, Ravenna, Italy*
[o]*Immunology Research Area, IRCCS Bambino Gesù Pediatric Hospital, Rome, Italy*

Received 15 December 2017; revised 21 March 2018; accepted 17 April 2018
Available online - 25 April 2018

ABSTRACT

Introduction: Determination of programmed death ligand 1 (PD-L1) expression defines eligibility for treatment with pembrolizumab in patients with advanced NSCLC. This study was designed to better define which value across core biopsy specimens from the same case more closely reflects the PD-L1 expression status on whole sections and how many core biopsy specimens are needed for confident classification of tumors in terms of PD-L1 expression.

Methods: We built tissue microarrays as surrogates of biopsies collecting five cores per case from 268 cases and compared PD-L1 staining results obtained by using the validated clone SP263 with the results obtained by using whole tumor sections.

Results: We found an overall positivity in 39% of cases at a cutoff of 1% and in 10% of cases at a cutoff of 50%. The maximum value across cores was associated with high concordance between cores and whole sections and the lowest number of false-negative cases overall. To reach high concordance with whole sections, four and three cores are necessary at cutoffs of 1% and 50%, respectively. Importantly, with 20% as the cutoff for core biopsy specimens, fewer than three cores showed high sensitivity and specificity in identifying cases with 50% or more of tumor cells positive for PD-L1 on whole sections. Specifically, for PD-L1

*Corresponding author.

Disclosure: The authors declare no conflict of interest.

Address for correspondence: Enrico Munari, MD, PhD, Department of Pathology, Sacro Cuore Don Calabria Hospital, Via Sempreboni 5, 37024 Negrar (VR), Italy. E-mail: enrico.munari@sacrocuore.it; enrico_munari@yahoo.it.

ISSN: 1556-0864
https://doi.org/10.1016/j.jtho.2018.04.017

expression values of 20% to 49% on cores, the probabilities of a tumor specimen expressing PD-L1 in at least 50% of cells on a whole section were 46% and 24% with one and two biopsy specimens, respectively.

Conclusions: An accurate definition of the criteria to determine the PD-L1 status of a given tumor may greatly help in selecting those patients who could benefit from anti–programmed cell death 1/PD-L1 treatment.

Keywords: PD-L1; Lung; Cancer; Heterogeneity; Biopsies

Introduction

Programmed cell death 1 (PD-1) is an inhibitory receptor that was originally identified in T lymphocytes. Upon interaction with its ligand programmed death ligand 1 (PD-L1), PD-1 delivers inhibitory signals that down-regulate T-cell function. Under physiological conditions, this interaction leads to peripheral T-cell tolerance, whereas in patients with cancer it may impair T-cell responses against tumor cells.[1–3] In this context, immunotherapy with checkpoint inhibitors that disrupt PD-1–PD-L1 interaction has proved highly effective in different tumor types, representing a true revolution in cancer therapy.[4–6]

Pembrolizumab is an anti–PD-1 humanized monoclonal antibody that was recently granted U.S. Food and Drug Administration approval after clinical trials conducted in patients with advanced lung adenocarcinoma or squamous cell carcinoma on the basis of PD-L1 expression on viable tumor cells, as assessed with a validated assay. Specifically, the KEYNOTE-010 trial demonstrated that pembrolizumab prolonged overall survival in previously treated patients whose tumors expressed PD-L1 in at least 1% of cells.[7] Moreover, the KEYNOTE-001 and the KEYNOTE-024 trials showed significantly longer progression-free survival and overall survival for previously untreated patients with tumors expressing PD-L1 in at least 50% of cells.[8,9] Therefore, the immunohistochemical evaluation of PD-L1 expression on tumor specimens has become an issue of major importance.

However, most studies focusing on the correlation between efficacy of PD-1/PD-L1 inhibitors and expression of PD-L1 on neoplastic cells do not take into consideration the type of material on which PD-L1 expression has been determined. In fact, such evaluation can be made on the basis of biopsy specimens or resection specimens. It is evident that PD-L1 expression in small tissue samples such as biopsy specimens might not be representative of the entire tumor specimen and can display divergent results because of cancer heterogeneity. In this regard, by using tissue microarrays (TMAs) as surrogates of biopsy specimens, we recently demonstrated that discordance between core biopsy specimens of a given tumor may occur in up to 20% and 7.9% of cases with use of cutoffs of 1% and 50% for PD-L1–positive cells, respectively.[10] To our knowledge, harmonization studies regarding determination of PD-L1 expression between core biopsy specimens and resection specimens are still lacking.

The present study was designed to better define which value across core biopsy specimens from the same case more closely reflects the actual PD-L1 expression status as determined on the resection specimen and the number of biopsy specimens needed at different cutoffs for consistent and reproducible PD-L1 quantification with respect to whole tumor sections, which are considered the criterion standard.

Material and Methods

Study Cohort

The study cohort consisted of consecutive patients with primary NSCLC who had undergone surgical resection at the Sacro Cuore Don Calabria Hospital of Negrar, Verona, Italy, between 2003 and 2017 and for whom slides and paraffin-embedded tissue blocks were available.

None of the patients received therapy before surgery.

Tumors were classified according to the 2015 WHO classification, and staging was done by using the TNM staging manual (seventh edition). Patient demographics and clinical data were retrieved from the digital archives.

The investigations were conducted according to the principles expressed in the Declaration of Helsinki.

TMA Construction

For every case, all hematoxylin and eosin–stained slides were reviewed for confirmation of diagnosis; one block was then selected for TMA construction. For each block, five cores with a diameter of 1 mm were obtained from diverse areas of the tumor and randomly numbered from 1 to 5. Overall, 12 TMAs were built.

Immunohistochemistry and Scoring

From each block (single cases and TMAs) 5-μm sections were cut and stained for PD-L1 (clone SP263, Ventana Medical Systems, Tucson, AZ) on an automated staining platform (Benchmark Ultra [Ventana Medical Systems]). An OptiView DAB IHC Detection Kit (Ventana Medical Systems) and an OptiView Amplification Kit (Ventana Medical Systems) were used according to the manufacturer's recommendations for visualization of the primary anti–PD-L1 antibody.

Stained sections were scanned with a Ventana iScan HT slide scanner (Ventana Medical Systems) and scored on the basis of percentage of tumor cells showing membranous positivity irrespective of staining intensities; a three-tiered system was then applied by using the following thresholds: less than 1%, 1% to 49%, and at least 50%.

PD-L1 evaluation was performed blindly by two pathologists who use clone SP263 in their clinical practice (E.M. and G.B.); discordant cases (cores and whole sections) were reevaluated by a third pathologist (G.R.).

Those cores showing a neoplastic component of at least 30% were considered adequate; therefore, cores with lower percentages of neoplastic component were excluded.

Macrophages were used as an internal control to validate the adequacy of the PD-L1 staining reaction.

Statistical Analysis

Statistical analysis was carried out with Stata software (StataCorp, College Station, TX) and R software (version 3.2.3, R Foundation for Statistical Computing, Vienna, Austria)[11]; χ^2 testing was used to analyze contingency tables and receiver operating characteristic curves were constructed to evaluate the predictive ability of core biopsy samples for PD-L1 status on whole sections. Cohen's κ was calculated for interrater agreement.

Results

Patient Characteristics

From the initial 271 patients, three were not considered for the analysis because of failure of TMA construction or staining. Overall, 268 patients were included in this study; of these, 190 (71%) were male and 78 (29%) were female; the median age was 70 years (range 41–87 years). The cancer of 183 of the 268 patients (68%) was diagnosed as adenocarcinoma, the cancer of 64 patients (24%) was diagnosed as squamous cell carcinoma, and 21 patients (8%) were determined to have other histotypes (15 patients had large cell carcinoma, four had adenosquamous carcinoma, and two had large cell neuroendocrine carcinoma). The median size of the tumors was 2.7 cm (range 0.8–21 cm).

Lymph node status was available for 246 patients: 172 patients were staged as N0, 35 as N1, 34 as N2, and five as N3.

PD-L1 Expression and Clinicopathological Features

Overall, PD-L1 expression value based on whole sections was negative (<1%) in 163 cases, between 1% and 49% in 79 cases, and at least 50% in 26 cases, with only one case showing homogenous positivity in 100% of cells.

We analyzed the associations between PD-L1 expression on whole sections and clinicopathological features (Table 1). We show that with a cutoff of 1%, PD-L1–positive tumors often corresponded to a higher stage. Notably, squamous cell carcinomas were more often positive than were adenocarcinomas.

Interobserver Variability and Heterogeneity of PD-L1 Expression

Regarding the evaluation of PD-L1 on whole sections, discrepancy between the two pathologists (E.M. and G.B.) was found in only six cases: in two cases there was lack of concordance at a cutoff of 1% and in four cases at a cutoff of 50% (Cohen's $\kappa = 0.98$ and 0.91 at cutoffs of 1% and 50%, respectively). After consensus, one case was defined as negative (<1%), one case was scored between 1% and 49%, and the other four cases were scored as at least 50% positive.

In the only case that showed 100% PD-L1 positivity on whole section, all five cores showed PD-L1 expression in at least 50% of cells. Of the 163 cases determined to be PD-L1–negative (<1% of stained cells) on whole sections, only five cases showed PD-L1 expression of at least 1% in some of the cores (one of five cores in two cases, three of five cores in one case, and two of four cores in two cases), with none showing positivity in more than 5% of tumor cells. Most of the cases had all five cores available for evaluation. In some cases, cores were lost during processing (n = 11) or contained less than 30% tumor cells (n = 72). Specifically, 215 cases had five cores, 31 had four cores, 16 had three cores, four had two cores, and two had one core available.

Regarding PD-L1 evaluation on the basis of cores, lack of concordance between the two pathologists (E.M. and G.B.) occurred in 21 cases (Cohen's $\kappa = 0.92$ and 0.82 at cutoffs of 1% and 50%, respectively), for which a third pathologist (G. R.) was required for consensus (Supplementary Table 1).

When all cores available for each case were considered, the concordance rates within the 104 cases that showed heterogeneity on whole sections were 93% at a cutoff of 1% and 88% at a cutoff of 50%. Notably, once all cases classified as negative had been eliminated, by definition there could not be any false-positive cases. Thus, discrepancies were due to the heterogeneous expression of PD-L1 in neoplastic cells (Fig. 1).

Definition of Best PD-L1 Value across Cores

We then compared the PD-L1 results obtained across tissue cores for each case to establish which value among the maximum, minimum, mean, and median

Table 1. Clinicopathological Characteristics of Patients in Relation with PD-L1 Expression on Whole Sections

		PD-L1 Expression ≥1%			PD-L1 Expression ≥50%		
Variable	Overall, n (%)	Negative, n (%)	Positive, n (%)	*p* Value[a]	Negative, n (%)	Positive, n (%)	*p* Value[a]
Patients	268	163 (61%)	105 (39%)		242 (90%)	26 (10%)	
Age							
<70 y	121 (45%)	76 (63%)	45 (37%)	*0.54*	111 (92%)	10 (8%)	*0.47*
≥70 y	147 (55%)	87 (59%)	60 (41%)		131 (89%)	16 (11%)	
Sex							
Male	190 (71%)	106 (56%)	84 (44%)	*0.01*	170 (89%)	20 (11%)	*0.48*
Female	78 (29%)	57 (73%)	21 (27%)		72 (92%)	6 (8%)	
Histologic type							
ADC	183 (68%)	120 (66%)	63 (34%)	*0.06*	164 (90%)	19 (10%)	*0.84*
SCC	64 (24%)	33 (52%)	31 (48%)		59 (92%)	5 (8%)	
Others	21 (8%)	10 (48%)	11 (52%)		19 (90%)	2 (10%)	
Diameter							
≤30 mm	164 (62%)	107 (65%)	57 (35%)	*0.09*[b]	151 (92%)	13 (8%)	*0.20*[b]
>30 mm	102 (38%)	56 (55%)	46 (45%)		89 (87%)	13 (13%)	
Unk	2	0	2		2	0	
N stage							
N0	172 (64%)	110 (64%)	62 (36%)	*0.14*[b]	158 (92%)	14 (8%)	*0.11*[b]
N1–N3	74 (28%)	40 (54%)	34 (46%)		63 (85%)	11 (15%)	
Unk	22 (8%)	13	9		21	1	

[a]Pearson chi-square test.
[b]For *p* value, unknown cases excluded from test.
PD-L1, programmed death ligand 1; ADC, adenocarcinoma; SCC, squamous cell carcinoma; Unk, unknown.

showed the highest rate of concordance with that obtained on the basis of whole sections (considered the criterion standard).

Overall, at a cutoff of 1%, we found concordance rates of 95.9%, 95.9%%, 86.5%, and 79.5% with the maximum, mean, median and minimum values, respectively. When a cutoff of 50% was used, we found concordance rates of 95.6%, 96.6%, 96.2%, and 94.4% with the maximum, mean, median, and minimum values, respectively. In general, the maximum and mean values appeared to better reflect PD-L1 expression on the whole sections. However, because the maximum value across cores showed the lowest number of false-negative results for cutoffs of both 1% and 50%, only this value was applied for subsequent analysis (Supplementary Table 2).

Definition of Minimum Number of Biopsy Specimens for Optimal Concordance with Whole Sections

Table 2 summarizes the results of sensitivity, specificity, and receiver operating characteristic analyses relative to the number of cores evaluated, with the results obtained with whole sections considered to be the reference for each case.

For cutoff of 1%, we found an area under the curve (AUC) greater than 0.9 with three core biopsy specimens; moreover, 0.91 sensitivity was reached when four cores were considered.

For cutoff of 50%, we found AUC and sensitivity values greater than 0.9 with three cores.

We noted that at cutoff of 50%, increasing numbers of core biopsy specimens corresponded to a decrease in specificity (from 0.98 with one core to 0.95 with five cores), reflecting the increasing number of false-positive cases. Overall, if we considered the cases with at least one core with PD-L1 expressed in at least 50% of cells, the false-positive rate was 3.7% (10 cases, all with at least four cores available) (Supplementary Table 3). Interestingly, if we considered cases with at least two cores with PD-L1 expressed in at least 50% of cells, the number of false-positive results dropped from 10 to three whereas the number of false-negative results did not change (two cases) (Supplementary Table 4).

Importantly, at a cutoff of 50%, if we considered fewer than three cores, sensitivity remained low (<90%). Therefore, we evaluated the predictive value of other PD-L1 cutoffs in core biopsy specimens, comparing the data with that in cases in which tumor cells expressed PD-L1 in at least 50% of cells on the whole sections. Thus, we assessed sensitivity, specificity, and AUC for cutoff values of 10%, 20%, 30%, and 40% related to the number of core biopsy specimens analyzed (Table 3). We found that at a cutoff of 20%, fewer than three cores were sufficient for identifying cases with at least 50% of tumor cells on whole sections testing positive for PD-L1 with a sensitivity and an AUC greater than 0.9. In particular, in the event of a single core with

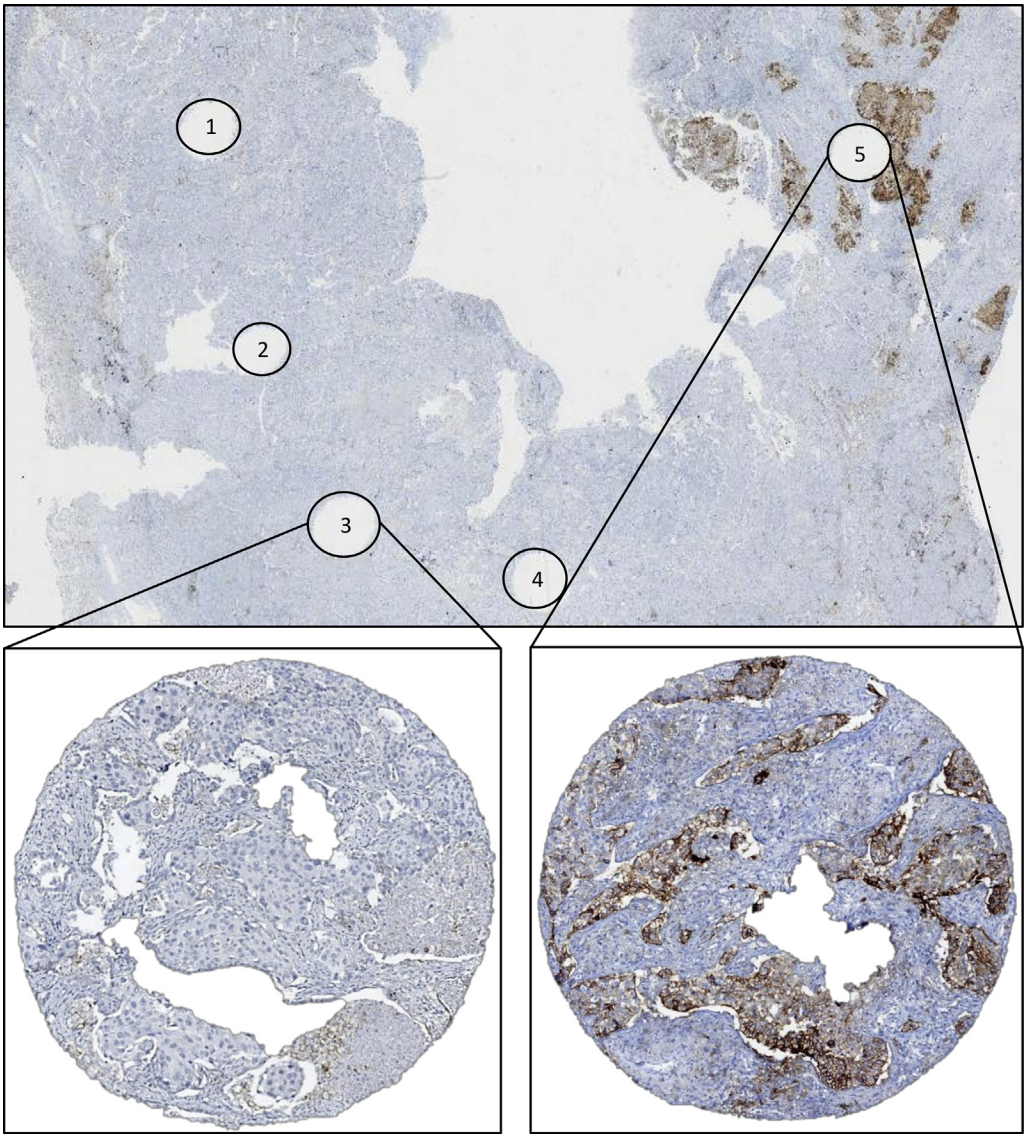

Figure 1. Representative image for heterogeneity of programmed death ligand 1 (PD-L1) expression. Whole section of squamous cell carcinoma of the lung stained with PD-L1 showing spots corresponding to tissue microarray cores numbered 1 to 5. Cores numbered 1 to 4 were sampled randomly in an area with a negative result when tested, whereas the result of testing of core 5 was positive in more than 50% of neoplastic cells. Overall, testing of the whole tumor yielded a result of PD-L1 expression in 20% of neoplastic cells.

20% to 49% of cells testing positive, the probability of having PD-L1 expressed in at least 50% of cells on whole section was 46%. In the event of two cores with an expression value of 20% to 49% in one or both, the probability of having PD-L1 expressed in at least 50% of cells on whole section was 24% (Table 4).

Discussion

Immunotherapy with PD-1/PD-L1 axis inhibitors represents a revolution in the field of oncology and a starting point for a new paradigm against cancer.[5] In this context, predicting which patients will respond to checkpoint inhibitor therapy is a major issue that so far has been based mainly on immunohistochemical evaluation of PD-L1 on tumor cells. Such evaluation may be performed both on resection specimens (when patients experience progression after tumor resection) and on small biopsy specimens (in the setting of advanced tumors at diagnosis).[9]

Recently, the U.S. Food and Drug Administration approved the anti–PD-1 pembrolizumab as a single agent for patients with tumors expressing PD-L1 in at least 50% of neoplastic cells for first-line therapy and as a second-line therapy for patients whose tumors express PD-L1 in at least 1% of cells.[7,8]

Moreover, although some studies have found significant correlation between expression of PD-L1 and response to therapy in different tumor types,[6,8,9,12,13]

Table 2. Sensitivity, Specificity, and ROC AUC according to Number of Available Cores by Cutoff

No. of Cores by PD-L1 Expression Cutoff	Cases, n	Sensitivity	Specificity	ROC AUC
PD-L1 ≥1%				
1	268	0.70	0.98	0.84
2	266	0.78	0.98	0.88
3	262	0.85	0.98	0.92
4	246	0.91	0.97	0.94
5	215	0.93	0.98	0.95
PD-L1 ≥50%				
1	268	0.77	0.98	0.87
2	266	0.88	0.98	0.93
3	262	0.92	0.98	0.95
4	246	0.95	0.96	0.96
5	215	0.95	0.95	0.95

ROC, receiver operating characteristic; AUC, area under the curve; PD-L1, programmed death ligand 1.

others have not, because responses have also been observed in patients whose tumors were classified as PD-L1–negative.[14,15] Such discrepancy could be due to different factors: on the one hand, there is the intrinsic and as-yet undiscovered biological complexity underpinning tumor-immunity interplay; on the other hand, it is reasonable to think that a percentage of tumors could be misclassified in terms of PD-L1 expression owing to factors such as expression heterogeneity,[10] different clones, and interpathologist variability.[16]

Therefore, it is of major importance to determine how many biopsy specimens should be obtained from tumors and how to consider the results of PD-L1 staining across them so as to maximize their reliability in predicting the true PD-L1 status of tumors and the probability of response to anti–PD-1/PD-L1 treatment.

In this study, we have addressed the question of which value of PD-L1 expression among maximum, mean, median, and minimum across tissue cores would best reflect the actual PD-L1 expression on the entire tumor specimen and how many biopsy specimens are necessary for optimal correlation. In addition, we have analyzed the predictive potential of different cutoffs for use with biopsy specimens in identifying cases with at least 50% of tumor cells expressing PD-L1 on whole tumor sections. To this end, we built TMAs as surrogates of diagnostic biopsies collecting five cores for each case from a total of 268 cases and compared the staining results with those obtained with whole tumor sections, which are considered the criterion standard. We found an overall positivity in 39% of cases for a cutoff of 1% and in 10% of cases for a cutoff of 50%. In general, the maximum value and the mean value across cores from each case showed the highest concordance rates with whole sections both at a cutoff of 1% (95.9% for both values) and at a cutoff of 50% (95.6% and 96.6%, respectively). Moreover, we found that the maximum value across cores correlated with the lowest number of the overall false-negative cases (four cases and two cases for cutoffs of 1% and 50%, respectively).

An important issue is definition of the minimum number of biopsy specimens necessary for optimal correspondence with whole tumor sections. Our results indicate that four and three core biopsy specimens are necessary to reach an AUC and sensitivity higher than 0.9 at cutoffs of 1% and 50%, respectively.

Importantly, with fewer than three core biopsy specimens, the sensitivity was not satisfactory (<0.9) at a cutoff of 50%. In this regard, we found that if we lowered the cutoff for cores to 20%, even just one or two cores were sufficient to identify cases with PD-L1 expressed in at least 50% of tumor cells on whole sections with high sensitivity and specificity.

In the clinical context, the main concern is to avoid missing patients who could benefit from an effective treatment and therefore keep the number of false-negative cases as low as possible; therefore, in the setting of a biopsy with at least 20% of tumor cells expressing PD-L1, it could be reasonable to consider the opportunity to repeat the bioptic procedure

Table 3. Sensitivity, Specificity, and ROC AUC of Different Cutoffs and Number of Cores in Predicting Cases with at Least 50% of Cells Positive for PD-L1 on Whole Sections

	Cutoffs of Cells Positive for PD-L1											
	10%			20%			30%			40%		
No. of Cores	SE	SP	AUC	SE	SP	AUC	SE	SP	AUC	SE	SP	AUC
1	.96	.91	.94	.96	.95	.96	.88	.96	.92	.81	.97	.89
2	.96	.88	.92	.96	.93	.94	.92	.95	.93	.88	.96	.92
3	.96	.85	.91	.96	.89	.93	.96	.92	.94	.92	.95	.93
4	1	.84	.92	.95	.89	.92	.95	.90	.93	.95	.95	.95
5	1	.85	.92	1	.89	.95	1	.90	.95	.95	.94	.94

ROC, receiver operating characteristic; AUC, area under the curve; PD-L1, programmed death ligand 1; SE, sensitivity; SP, specificity.

Table 4. Concordance between Cores with Less than 50% of Cells Expressing PD-L1 and Whole Sections, with 20% Used as Cutoff for Cores

	Whole Sections	
No. of cores	PD-L1 Expression <50%	PD-L1 Expression ≥50%
1		
<20%	231	1
20%-49%	6	5
2		
<20%	223	1
20%-49%	13	4

PD-L1, programmed death ligand 1.

before excluding a patient for first-line therapy with pembrolizumab.

Because treatment with checkpoint inhibitors (including PD-1/PD-L1 inhibitors) may cause some toxicity, avoiding false-positive cases is also of major importance. In this regard, it should be noted that in our cohort at a cutoff of 50%, when at least two cores were positive, the risk of false-positive cases dropped to 1.2%.

So far, only a few studies have addressed the issue of PD-L1 heterogeneity and its potential role in tumor misclassification, given the discrepancies between biopsy specimens and resection specimens.[17–19] Moreover, no attempts have been made to harmonize biopsy specimens and resection specimens with regard to determination of PD-L1 expression.

Kitazono et al. evaluated PD-L1 expression by using a polyclonal antibody (4059, ProSci, Poway, CA) on 79 diagnostic biopsy specimens and corresponding surgical specimens; they found concordance rates of 92.4% and 83.5% for cutoffs of 1% and 50%, respectively.[19] In another similar work, by evaluating PD-L1 using the SP142 clone on 160 surgically resected samples and paired diagnostic biopsy specimens, Ilie et al. found a concordance rate of 81% with use of a cutoff of 1% for tumor cells.[18]

These two studies[18,19] differ from our study in that the authors used diagnostic biopsy specimens; however, the percentage of neoplastic specimens out of total biopsy specimens is not clear in either of the studies. Ilie et al.[18] found a trend toward a significant difference between the average number of diagnostic biopsy specimens in discordant versus concordant cases (3.4 versus 6.8).

Other works used TMAs as surrogate of diagnostic biopsy specimens. Gniadek et al. evaluated PD-L1 expression on TMAs from 150 cases of NSCLC by using SP142 antibody and an Abcam detection kit (Abcam, Cambridge, MA); they found discrepancies among cores in a total of 28 of 71 positive cases (40%).[17] Unlike us, however, they used whole sections only in cases that showed discordance within cores.

Recently, Li et al. evaluated PD-L1 expression on 190 resected NSCLC and matched TMA cores by using the Dako 22C3 clone (Dako, Carpinteria, CA). At a cutoff of 1% these authors found 37% of cases to be positive, with a discordance rate of 13.2%, whereas at a cutoff of 50% positivity was found in 11% of cases, with a discordance rate of 6.8%.[20] Notably, these authors built their TMA by using single cores with a diameter of 2 mm for each case; instead, we used multiple smaller cylinders (1 mm) to have a more comprehensive picture for each case and to allow better computation thanks to a higher number of cores. In addition, even with consideration of our results on only a single core (and even though the core was smaller), our data are in line with those of these authors in terms of discordance rates (13% and 4.2% for cutoffs of 1% and 50%, respectively).

In relation to general positivity at cutoffs of 1% and 50%, the results reported by Li et al.[20] are similar to ours even though two different clones were used (22C3 by Li et al. and SP263 by us).

We realize that a limitation of our study lies in its retrospective nature and the use of TMAs as surrogates of tumor biopsy specimens. In fact, in routine diagnostics, not all biopsy specimens actually correspond to neoplastic tissue; therefore, the applicability of our approach is possible if only neoplastic biopsy specimens are considered. However, we believe that such approach is informative, as the neoplastic component present in a transbronchial biopsy that we encounter in our clinical practice is similar in terms of quantity to the neoplastic component present in each tissue core.

Another point is related to the issue of interobserver variability, as was recently demonstrated in an article by Brunnstrom et al.[16]; in this regard, all cores and whole sections were evaluated blindly by two pathologists who use SP263 in their clinical practice and discordant cases were reevaluated by a third pathologist for consensus.

In conclusion we have demonstrated that when evaluating multiple biopsy specimens for PD-L1 assessment, the maximum value across cores should be considered. Moreover, to reach high concordance with whole sections, four and three core biopsy specimens are necessary at cutoffs of 1% and 50%, respectively. In addition, in the first-line setting, when at least two cores result in positive at a cutoff of 50%, the rate of false-positive cases drops to 1.2%. Importantly, in the event of fewer than three cores, of which at least one has at least 20% PD-L1–positive cells, the probability of a tumor with a level of PD-L1 expression of at least 50% on the whole section is high, with sensitivity and AUC greater than 0.9. In our view, according to our results, pathologists should describe in their report the number

of cores analyzed and the percentage of cells positive for PD-L1 in each one, thus allowing clinicians to make a more confident decision with regard to patient selection for therapy.

Acknowledgments

This work was partially supported by the Italian Association for Cancer Research (AIRC) (IG 2014 15283 to Dr. Lorenzo Moretta).

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org and at https://doi.org/10.1016/j.jtho.2018.04.017.

References

1. Ishida Y, Agata Y, Shibahara K, Honjo T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J*. 1992;11:3887–3895.
2. Nishimura H, Nose M, Hiai H, Minato N, Honjo T. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity*. 1999;11:141–151.
3. Pesce S, Greppi M, Tabellini G, et al. Identification of a subset of human natural killer cells expressing high levels of programmed death 1: a phenotypic and functional characterization. *J Allergy Clin Immunol*. 2017;139:335–346.e333.
4. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12:252–264.
5. Brahmer JR, Pardoll DM. Immune checkpoint inhibitors: making immunotherapy a reality for the treatment of lung cancer. *Cancer Immunol Res*. 2013;1:85–91.
6. 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;366:2443–2454.
7. 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;387:1540–1550.
8. 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;375:1823–1833.
9. Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372:2018–2028.
10. Munari E, Zamboni G, Marconi M, et al. PD-L1 expression heterogeneity in non-small cell lung cancer: evaluation of small biopsies reliability. *Oncotarget*. 2017;8:90123–90131.
11. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed May 11, 2018.
12. Brahmer JR, Tykodi SS, Chow LQ, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med*. 2012;366:2455–2465.
13. 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;20:5064–5074.
14. 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;373:1627–1639.
15. Brahmer JR, Hammers H, Lipson EJ. Nivolumab: targeting PD-1 to bolster antitumor immunity. *Future Oncol*. 2015;11:1307–1326.
16. Brunnstrom H, Johansson A, Westbom-Fremer S, et al. PD-L1 immunohistochemistry in clinical diagnostics of lung cancer: inter-pathologist variability is higher than assay variability. *Mod Pathol*. 2017;30:1411–1421.
17. Gniadek TJ, Li QK, Tully E, Chatterjee S, Nimmagadda S, Gabrielson E. Heterogeneous expression of PD-L1 in pulmonary squamous cell carcinoma and adenocarcinoma: implications for assessment by small biopsy. *Mod Pathol*. 2017;30:530–538.
18. Ilie M, Long-Mira E, Bence C, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. *Ann Oncol*. 2016;27:147–153.
19. Kitazono S, Fujiwara Y, Tsuta K, et al. Reliability of small biopsy samples compared with resected specimens for the determination of programmed death-ligand 1 expression in non-small-cell lung cancer. *Clin Lung Cancer*. 2015;16:385–390.
20. 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.