Programmed Death-Ligand 1 Immunohistochemistry in Lung Cancer

In what state is this art?

Keith M. Kerr, MBChB, FRCPath,* Ming-Sound Tsao, MD, PhD,† Andrew G. Nicholson, DM, FRCPath,‡ Yasushi Yatabe, MD, PhD,§ Ignacio I. Wistuba, MD, PhD,¶ and Fred R. Hirsch, MD, PhD,**

On behalf of the IASLC Pathology Committee

Abstract: Therapeutic antibodies to programmed death receptor 1 (PD-1) and its ligand PD-L1 show promising clinical results. Anti-PD-L1 immunohistochemistry (IHC) may be a biomarker to select patients more likely to respond to these treatments. However, the development of at least four different therapeutic agents, each with a different anti-PD-L1 IHC assay, has raised concerns among pathologists and oncologists alike. This article reviews existing data on the IHC biomarker aspects of studies using these drugs in non–small-cell lung cancer (NSCLC) and considers the challenges ahead, should these drug/IHC assay combinations reach routine practice. For each of the known biomarker assays in development, there is a different monoclonal IHC antibody clone, produced by one of two diagnostics companies. Each test requires proprietary staining platforms and uses different definitions of a “positive” test for PD-L1 expression, on tumor cells and, in one test, also on tumor infiltrating immune cells. There are still considerable gaps in our knowledge of the technical aspects of these tests, and of the biological implications and associations of PD-L1 expression in NSCLC, considering heterogeneity of expression, dynamic changes in expression, and prognostic implications among other factors. The International Association for the Study of Lung Cancer Pathology Committee raises the prospect of trying not only to harmonize and standardize testing for PD-L1 by IHC, at least at a technical level, but also, ideally, as a predictive marker, to facilitate availability of this test and a promising treatment for patients with NSCLC.

Key Words: Immune check-point inhibitors, PD-1, PD-L1, Immunohistochemistry, Biomarker assay.

(J Thorac Oncol. 2015;10: 985–989)

IMMUNE CHECKPOINT INHIBITION: A PROMISING THERAPEUTIC STRATEGY FOR LUNG CANCER

In the search for effective therapies in patients with lung cancer, immune checkpoint inhibitory approaches have shown considerable promise. A number of ligand–receptor interactions, including PD-1/PD-L1 and B7/CTLA-4, seem to switch off the immune response in lung cancer, a tumor that in general has a high rate of somatic mutations, which may make such tumors more immunogenic. Much of this therapeutic focus in lung cancer, particularly in non–small-cell lung cancer (NSCLC), has been on interrupting the interaction of programmed death receptor-1 (PD-1) and its ligand (PD-L1) between tumor cells and immune effectors cells, using monoclonal antibodies against PD-L1 or PD-1. In this era of personalized medicine using targeted biological agents, biomarkers predictive of response to therapy are central to treatment decision making.

AVAILABLE THERAPIES AND BIOMARKERS

There are a number of therapeutic anti-PD-L1 (e.g., MPDL3280A [Roche, Basel, Switzerland] and MEDI-4736 [Astra Zeneca, London, UK]) or anti-PD-1 (nivolumab [Bristol Myers Squibb, New York, NY]) and pembrolizumab [Merck, Kenilworth, NJ]) agents at various stages of development, and the favored biomarker seems to be the expression of PD-L1 assessed by immunohistochemistry (IHC; Fig. 1). There are limited data currently available, for these therapeutic agents, in lung cancer, in particular in patients with advanced NSCLC. Different approaches have been taken to PD-L1 IHC assessment, using different diagnostic antibodies to assess PD-L1 expression, different technical staining platforms, and different definitions of a “positive” predictive IHC stain. In some cases, expression of PD-L1 on immune effector cells as opposed to, or in combination with, expression in tumor cell, has been chosen as the biomarker.
PROBLEMATIC ISSUES WITH EXISTING DATA

Some of the essential findings so far reported are presented in Table 1.7–20 Data are limited and most remain unpublished at the time of writing. Depending on definitions, positivity rates for PD-L1 range from 13% to 70%, and correlation between biomarker positivity and treatment response rates vary from 13% to 83% depending upon the biomarker-defined cohort and therapy used. Most studies also report significant response rates (3–20%) in PD-L1 IHC negative cases. Most of the studies assess PD-L1 expression in tumor cells and regard membrane staining as most significant. There is variable interpretation of the intensity and distribution of staining and variable definition of a positive PD-L1 stain ranging from staining of ≥1% to ≥50% of cells assessed. In some cases, the test requires at least 100 tumor cells to be assessed.

Biomarker Positivity and Response

The value of the chosen biomarker seems to vary in terms of predicting a response to therapy, and in some cases this also seems to depend on which line of therapy for which the immune checkpoint inhibitory agent is given (Table 1). The biomarker test may not represent the true PD-L1 status of the tumor, either because of heterogeneity of expression and sampling error, or because the test sample predates earlier lines of therapy (see below). In general, however, there is a higher response rate in the PD-L1 positive population compared with the PD-L1 negative group of patients, although in some studies this difference is not significant. The presence of patients who respond to therapy, in the PD-L1 negative cohort, calls into question the value of PD-L1 IHC as a predictive biomarker to select a patient subgroup for therapy.

Biomarker Thresholds

Determining the threshold that defines a positive, predictive test is a difficult issue. Thresholds may be predetermined, before outcome data are known, or as a more useful approach, the response data may be used to indicate the threshold that gives best discrimination between responders and nonresponders, or between patients who do or do not derive significant survival benefit from the therapy. It has, however, been noted that traditional response evaluation criteria in solid tumors for assessing tumor response may not be best suited to assessing clinically significant responses to immune checkpoint inhibitor therapy, at least in a small proportion of the cases. There is then a potential trade-off between improving upon the response rates seen in an unselected treated population, the acceptability of this response rate in an unselected population versus that seen with standard of care treatment, and any considerations to maximize the population eligible for treatment. In addition, to date, response (overall response rate) alone does not seem to be the best way to evaluate the benefit of immunotherapy; this is probably better captured by progression-free or overall survival data. Finally, if very low staining thresholds such as 1% or even 5% of cells are chosen, there is a greater risk that scoring will be inconsistent and is more likely to reflect inaccurately the patient’s tumor burden overall, because of heterogeneity.

Heterogeneity and Prior Therapy

Limited data suggest that PD-L1 expression is heterogeneous, reflected in low thresholds being used to define positive staining. Little is understood regarding the relationship between PD-L1 expression in the primary tumor and any metastases. Earlier lines of chemotherapy or targeted therapy may well induce PD-L1 expression, consequently PD-L1 expression in the original “chemo-naive” diagnostic

FIGURE 1. Programmed death receptor-1 with its ligand (PD-L1) immunostaining performed using the E1LN3N clone anti-PD-L1 from Cell Signaling Technology (Boston) with standard detection techniques. A, Squamous cell carcinoma showing a strong, uniform positive reaction in tumor cells. B, Despite being negative in tumor cells in the center of the image, there is a positive reaction in macrophages and other immune cells in the tumor stroma. C, Most alveolar macrophages are positive for PD-L1. D, This adenocarcinoma is negative for PD-L1. It should be noted that this immunohistochemistry clone was not used for PD-L1 detection in any of the trials discussed in this review.
sample may not represent the status of the tumor at the time that an immune checkpoint inhibitor therapy is introduced. This dynamic property may explain why biomarker data have not necessarily predicted responses when some of these drugs are given in second or later lines, reflecting response rates of 10–20% while the biomarker was negative in the chemo-naive sample. For most existing data, it is impossible to adequately relate any lymphoid or other diagnostic samples are so small and disaggregated that it is impossible to adequately relate any lymphoid or other immune effector cells to the tumor present. For this reason, a biomarker test based upon tumor infiltrating immune cells would almost certainly rule out lymph node biopsy samples and possibly all cytology specimens.

TABLE 1. Summary of Published Findings for PD-L1 Immunohistochemistry in Therapeutic Trials

<table>
<thead>
<tr>
<th>Drug</th>
<th>Biomarker Antibody</th>
<th>Rx Line</th>
<th>Definition of &quot;Positive&quot;* (%)</th>
<th>N Positive (%)</th>
<th>Positive Predictive Outcome</th>
<th>ORR % IHC pos. Cases</th>
<th>ORR % IHC neg. Cases</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nivolumab</td>
<td>Dako 28-8</td>
<td>1st</td>
<td>≥5 in &gt;100 cells</td>
<td>59</td>
<td>Yes</td>
<td>31</td>
<td>10</td>
<td>7,8'</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>Dako 28-8</td>
<td>≥2nd</td>
<td>≥5, ≥1</td>
<td>49, 56</td>
<td>No</td>
<td>15, 13</td>
<td>14, 17</td>
<td>9,10</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>Dako 28-8</td>
<td>1st</td>
<td>≥5 in &gt;100 cells</td>
<td>42</td>
<td>No</td>
<td>19</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>Dako 28-8</td>
<td>≥2nd</td>
<td>≥5</td>
<td>33'</td>
<td>Yes</td>
<td>24</td>
<td>14</td>
<td>12'/</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>5H1</td>
<td>≥2nd</td>
<td>≥5, also studied</td>
<td>67</td>
<td>Yes</td>
<td>No data</td>
<td>No data</td>
<td>13</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>Dako 22C3</td>
<td>Any</td>
<td>“Strong” ≥50, “Weak” 1–49</td>
<td>25, 70</td>
<td>Yes, Yes</td>
<td>37, 17</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>Dako 22C3</td>
<td>1st</td>
<td>≥50, ≥1</td>
<td>?</td>
<td>Yes</td>
<td>47, 26</td>
<td>?</td>
<td>15</td>
</tr>
<tr>
<td>MPDL3280A</td>
<td>Roche Ventana, SP142</td>
<td>≥2nd</td>
<td>≥10, ≥25, ≥1 TIICs</td>
<td>13, 28, 56</td>
<td>Yes</td>
<td>83, 46, 31</td>
<td>18, 18, 20</td>
<td>16–18</td>
</tr>
<tr>
<td>MEDI-4736</td>
<td>Roche Ventana, SP263</td>
<td>≥2nd</td>
<td>Data not available</td>
<td>41</td>
<td>Yes</td>
<td>25</td>
<td>3</td>
<td>19,20</td>
</tr>
</tbody>
</table>

*Expression in tumor cells unless otherwise stated.

TEST REPRODUCIBILITY AND EPITOPE STABILITY

Inevitably, whenever an IHC-based biomarker is considered, questions arise about the reproducibility of the test, not only in technical terms for producing the staining but also in interpretation of the test by pathologists. Furthermore, how stable are the epitopes detected by the various antibodies, which raises issues about the use of stored, pre-cut sections. Preanalytical issues such as tissue fixation and processing can have a major impact on the outcomes of immunohistochemical reactions, and how these might affect the different reported PD-L1 IHC tests is not known.

MULTIPLE DRUGS AND MULTIPLE BIOMARKER ASSAYS

Notwithstanding the difficulties there may be in delivering a robust biomarker assessment for PD-L1 IHC, how shall the pathology community handle the prospect of multiple different tests, ostensibly measuring the same biomarker that determine the prescription of several different therapeutic agents targeting the same molecular mechanism? Our experience of the development of companion IHC diagnostics suggests that these biomarkers may become available only in the form of a prepackaged test kit of reagents. The benefits of not only such standardization, but also the associated costs, are well understood. However, these kits normally mandate the use of a company-specific automated staining platform. Many pathology departments may be constrained by available technology and may not be able to carry out a required test, which may lead to low screening rates and patients missing out on targeted therapies if access to the drug is predicated on a specific, and more expensive, test requiring company-specific staining platforms.

Copyright © 2015 by the International Association for the Study of Lung Cancer
The potential existence of multiple different tests with different scoring systems will also be a recipe for confusion, among oncologists and pathologists alike. This could be compounded by the need for different scoring systems in different tumor types, using the same antibody test. Experience tells us that where there is confusion, or complication in test assessment, attitudes toward the test are changed. This has an impact on both the willingness of pathologists to provide the test and probably the quality of the data delivered from assessing the test staining. In large institutions where several of these drugs may be used, there could be significant logistical issues in terms of ensuring that the appropriate test is carried out. Unless an institution focuses on only one drug, or there is a level of communication between oncologist and pathologist that is, to date, unprecedented, it will be impossible to undertake reflex testing for PD-L1 expression. There is a considerable danger that these issues could have a significant impact on an institution’s interest or ability to use these therapies.

Is there any possibility for test harmonization? The traditional model, prospectively proving a specific biomarker in a clinical trial, is well understood. Scientific rigor determines that any deviation from what was shown in the trial to be effective may lead to different and misleading results, a situation which is unacceptable when considering companion diagnostics. Is there an opportunity to somehow determine equivalence between a number of available diagnostic anti-PD-L1 IHC antibodies? And if equivalent staining performance can be demonstrated, what are the chances of developing a standardized scoring system, across test platforms, that may be shown to be predictive for response to a number of these promising therapies? A concerted, multicenter, international effort could provide a mechanism to do such work, but this would require collaboration by both pharmaceutical and diagnostic companies with academic pathologists, to facilitate access to both the diagnostic tests and trial outcome data. Such a study could eventually also be extended to identify whether alternative biomarkers could either replace, or be combined with, PD-L1 as a predictive classifier, recognizing the complexity of the immune system.

QUESTIONS TO BE ADDRESSED

PD-L1 IHC seems to be an encouraging predictive biomarker for anti-PD-L1/PD-1 therapy in NSCLC, but given our current state of knowledge and understanding, the pursuit of a PD-L1 IHC assay as a “companion diagnostic” raises many issues. Questions still remain to be addressed regarding the biology of PD-L1 expression, including heterogeneity, correlations with stage of disease, ethnic associations, demographic characteristics, impact of prior lines of therapy, and the associated co-medications including steroids, mutation status of the tumor, and any prognostic effects. Our understanding of the prognostic significance of PD-L1 expression is not clear, yet this will have import interpretation of biomarker-based treatment outcomes. These are important issues, if patients are to be appropriately selected for, or denied, a potentially effective treatment. In considering this effectiveness, most data refer to overall response rate, while progression-free and overall survival may be more valid measures, and we have yet to see prospective, randomized phase II or III trials comparing these immunotherapies with standard chemotherapy. Regarding the assay itself, the influences of preanalytical variables, applicability across different staining platforms, usage on different sample types (large tumor samples, small biopsies, and cytology), intralaboratory and interlaboratory reproducibility, intraobserver and interobserver variability, and epitope stability in stored materials are matters for further study.

PROPOSAL FOR MULTICENTRE INTERNATIONAL STANDARDIZATION PROJECT

A multicenter, international standardization effort could address many of these questions and help develop one “standardized” assay, for each of this family of drugs that comes into clinical use and analyze additional immunotherapy-related predictive markers. Of course, these therapeutic agents are at different stages of development, and there is no guarantee that all of them will reach the market, but if more than one does, then the issues described above become highly relevant, assuming the biomarker is required to select patients. Commercial considerations will undoubtedly be an issue in terms of what can be done. It will be a disservice to our patients, however, if the complications discussed in this review have impact upon the availability of a valuable treatment.

REFERENCES

PD-L1 IHC in Lung Cancer


APPENDIX

International Association for the Study of Lung Cancer (IASLC) Pathology Committee members: Chair—AG Nicholson, Royal Brompton Hospital, London, United Kingdom. Biomarker subgroup—KM Kerr (IASLC Board Liaison), Aberdeen University Medical School, Aberdeen, United Kingdom; Y Yatabe, Aichi Cancer Centre, Nagoya, Japan; MS Tsao, Princess Margaret Hospital, Toronto, Canada; IIWistuba, MD Andersen cancer Center, Houston, Texas; and FR Hirsch (CEO IASLC), University of Colorado Cancer Center, Denver, Colorado. Members—MB Beasley, Mount Sinai Hospital, New York; E Brambilla, CHU Albert Michallon, Grenoble, France; J Botling, Uppsala University, Sweden; LR Chirieac, Brigham and Women’s Hospital, Boston, Massachusetts; S Dacic, University of Pittsburgh Medical Center, Pennsylvania; G Pelosi, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; Y Ishikawa, Japanese Foundation for Cancer Research, Tokyo, Japan; N Lertprasertsook, Chiang Mai University, Chiang Mai, Thailand; A Moreira, Memorial Sloan Kettering Cancer Center, New York, New York; M Noguchi, University of Tsukuba, Tsukuba, Japan; I Petersen, Jena University Hospital, Jena, Germany; E Thunnissen, VUMC, Amsterdam, Netherlands; KF To, Chinese University of Hong Kong, Hong Kong; and WD Travis, Memorial Sloan Kettering Cancer Center, New York, New York.