

Making the best of PD-L1 IHC testing

Anne Paxton

July 2016—When Keith Kerr, MB ChB, describes the ideal biomarker, he isn't hesitant about what pathologists and clinicians need. "Ideally, the biomarker would always be correct. It would be easy and practical to measure. It would either be present or absent, with no gray zone or doubt. The biomarker itself would be a stable and functionally unique factor related to the system being studied. And it would, of course, be 100 percent predictive."

PD-L1 testing can't reach that standard. But how far short of the ideal does PD-L1 testing fall? Dr. Kerr, consultant pathologist at Scotland's Aberdeen Royal Infirmary and a professor of pulmonary pathology at the Aberdeen University Medical School, discussed PD-L1 testing's key weaknesses and how pathologists can grapple with them, during a May webinar produced by CAP TODAY in collaboration with Dako. He sees immunotherapy as intensifying pressure on pathologists to produce reliable results—even as the nature of PD-L1 expression, the variability of scoring, and the lack of consensus standards make reliability challenging.

PD-L1 testing is imperfect, the factors surrounding evaluation of results are tricky issues, and pathologists are unavoidably treading in difficult territory, Dr. Kerr said. "But the oncology world is watching. It's really incumbent upon the pathology community to get this right, or to get it as right as we can. If we cannot deliver good outcomes with this particular test, I think we will be doing immunohistochemistry a disservice."

The stakes of the diagnostics that guide patient choices about therapy are high, said webinar co-presenter Annika Eklund, PhD, global product manager for companion diagnostics at Agilent Technologies (Dako is an Agilent company). Agilent developed the first two FDA-approved PD-L1 assays: a companion diagnostic (PD-L1 IHC 22C3 pharmDx) for Merck's drug Keytruda (pembrolizumab) and a complementary diagnostic (PD-L1 IHC 28-8 pharmDx) for Bristol-Myers Squibb's drug Opdivo (nivolumab).

The value of PD-L1 immunohistochemistry—for example, in non-small cell lung cancer—is that immunotherapy is more likely to help a PD-L1-positive patient's T cells to stay active and attack the tumor, while a PD-L1-negative patient is less likely to be helped in that way, Dr. Eklund said. "PD-L1 testing is valuable, too, to identify which patients are most likely to respond to therapy."

The CheckMate 057 study of nivolumab versus the chemotherapy docetaxel showed that nivolumab patients with a PD-L1 expression of more than one percent had 17.1 months median overall survival, versus nine months for docetaxel patients. Previous research has indicated that patients' PD-L1 expression can be dynamic, Dr. Eklund added. "However, experience in NSCLC supports that either fresh or archival tissue can be assessed by PD-L1 IHC."

To develop standardized and evidence-based PD-L1 IHC, the company starts by integrating the primary antibody into a standardized solution, optimizing reagents for antigen retrieval, and developing controls. It then optimizes the protocol with software and automation, Dr. Eklund said. "We develop interpretation and scoring guidance to identify responding patients for the unique drug, and this complete standardized solution is validated in clinical trials to provide reliable test performance."

Users should not optimize these tests on their own, she warned, as any deviation from the validated practice could be harmful to patients.



So far, in lung cancer, Dr. Kerr said, the field of immunodiagnostics has tackled mainly the “low-hanging fruit” among potential biomarkers. “Our biomarker experiences are largely around *EGFR* mutation and *ALK* translocation, both addictive oncogenes that have very particular biological features and are the main, if not the sole, driver of the tumor. Yet the best response rates we’ve seen for selected drugs are 60 to 70 percent, not 100 percent. And we all know the testing is far from being foolproof.”

The new world of immunotherapy is complicated, he said, with many different cell types and many, many different molecular factors, soluble factors, and cell-to-cell interactions. “And the therapy we are thinking about and the biomarker we are thinking about are only one single factor in this very complex mix of things going on.”

As one of the most antigenic of solid tumors, lung cancers will have a high number of neoantigens and non-self antigens expressed on the tumor cells. “It’s very likely that a competent immune system will eliminate abnormal clones of neoplastic or near-neoplastic cells before they ever cause any clinically relevant cancer. So somehow, for a tumor to become clinically evident, it must escape the attention of the immune system,” Dr. Kerr said.

One likely escape route is through the actions of immune checkpoints, complicated systems of receptors that switch immune sector cells on or off when ligands are encountered. The interaction between PD-1 on the surface of T effector cells and its ligand PD-L1 is one of these possibly inhibitory interactions. In physiological terms, “We believe that these interactions are probably responsible for preventing autoimmune responses, for switching off our immune response to self antigens. It seems likely that some cancer cells manage to hijack this system in order to evade the immune system,” Dr. Kerr said.

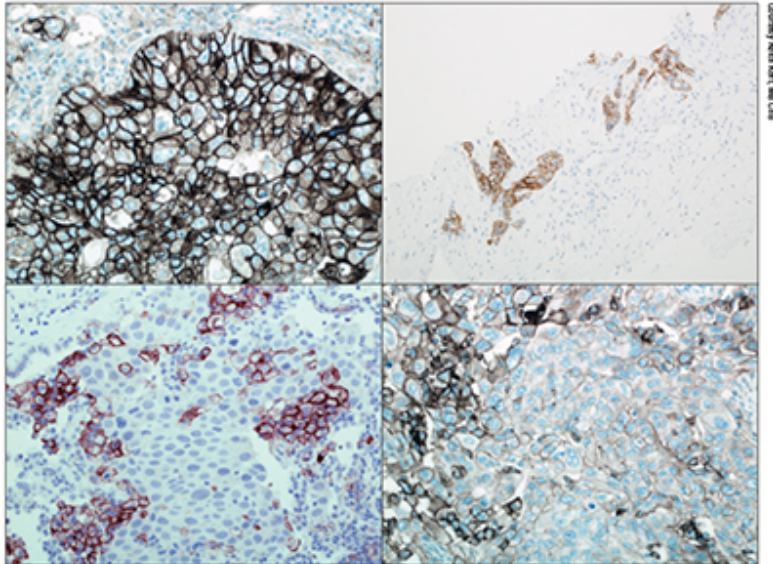
It has been fairly consistently observed that PD-L1 expression, at least with non-small cell lung cancer, is a poor prognostic feature, indicating that the presence of PD-L1 somehow manages to switch off an immune system that otherwise might improve a patient’s prospects. Targeting and inhibiting the checkpoints might therefore be a sound approach to therapy, Dr. Kerr said. “The multiple interactions of PD-1 and PD-L1 with lymphocytes, T effector cells, and a whole range of other cell types, including tumor cells, might manage to convince an immune system that has been switched off to become reactivated and to attack the tumor.”

In response to an audience question, Dr. Kerr said that over the decades, “there is hardly a single consistent IHC biomarker that has been shown to be either a positive or negative prognostic in lung cancer.” He suspects this pattern has much to do with the enormous heterogeneity in the way immunohistochemistry is performed, using different clones and scoring standards, among other differences.

Several clinical studies in NSCLC have found higher response rates to anti-PD-1 or anti-PD-L1 drugs in patients with higher PD-L1 expression. “And we are now seeing that this effect and response rate in relation to PD-L1 expression are beginning to translate into progression-free survival and particularly into overall survival in those patients.”

Percentages and cutoffs in PD-L1 evaluation, however, vary in their significance.

A phase two trial for the 22C3 clone developed as a companion diagnostic for Merck’s pembrolizumab showed superior overall survival for previously treated NSCLC patients with a PD-L1 score greater than 50 percent, he noted. But a phase three trial made the additional observation that overall survival in the 1-49 percent group was better with pembrolizumab than the docetaxel group.



Non-small cell lung cancer samples stained with the 28-8 Dako IHC assay. The top two images show widespread tumor cell membrane staining. Bottom right also shows membrane staining but note the variability, with significant numbers of negative tumor cells. Bottom left shows membrane staining in macrophages, but no staining in tumor cells.

The POPLAR trial, recently published, studied the biomarker for atezolizumab, finding an improving positive ratio in favor of atezolizumab response as the degree of either immune cell or tumor cell staining increases. However, Dr. Kerr noted, somewhat different numbers are used to define the cutoffs for immune cell and tumor cell scores, particularly at the high end of the scale. About one-third of the patients were deemed positive because of immune cell staining only.

In the CheckMate trials of nivolumab, using an anti-PD-L1 IHC assay, based on a clone known as 28-8, “we have a threshold of definition of positivity of one percent and membrane staining of tumor cells only.” But, Dr. Kerr added, “As definitions of positivity change—above one percent, above five percent, above 10 percent—the chance of patient benefit improves, though in all instances this is better than the docetaxel control arm.”

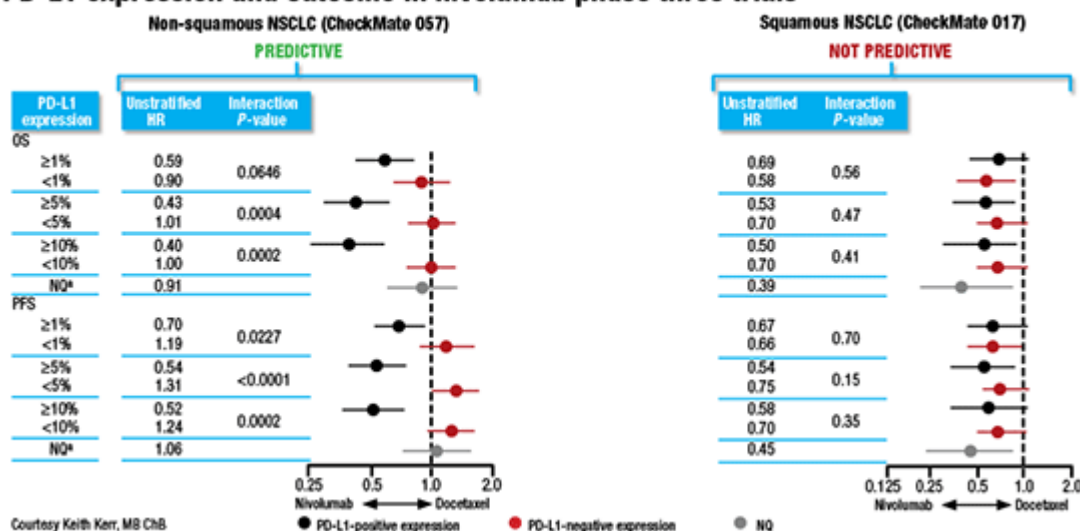
Although the numbers in the trial are small, “this rather implies that most of the effect, perhaps like pembrolizumab, is being driven by the patients who are highly expressing PD-L1.” (This pattern is for non-squamous NSCLC. Use of PD-L1 IHC for squamous NSCLC is not required, he noted, for nivolumab because the biomarker is not predictive in this setting.)

Some alternative biomarkers are showing promise, Dr. Kerr reported. Immune gene signatures, which measure expression genes by messenger RNA extracted from tumor samples, are one example. Far more data are needed, in part because mRNA is more complicated and may be more fragile to extract in adequate quantity and quality from diagnostic samples. “But it’s certainly an interesting alternative for the future,” he said.

Immune cells themselves, in addition to PD-L1 expression, are also being considered as potential biomarkers. “For example, looking at the presence or absence of particular types of immune cells and where the infiltrate might be in the tumor: Is it amongst the tumor cells or is it only adjacent to the tumor? We know that these factors are prognostic. Are they also predictive of response to therapy? Again, we need more data to draw reasonable conclusions.”

Interferon gamma is also an important regulator of PD-L1 expression, he said. “So more studies are now looking at the presence or absence of other immune checkpoints and interferon gamma expression, possibly as an adjunct to the PD-L1 expression as a biomarker test for these therapies.”

PD-L1 expression and outcome in nivolumab phase three trials



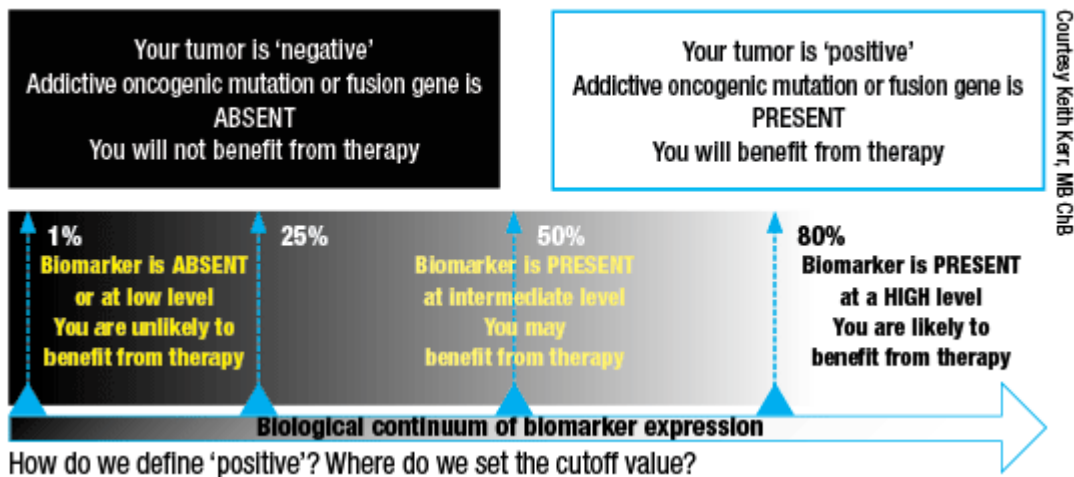
The concept of mutation burden is a fourth possibility, Dr. Kerr said. One study looked at mutation burden as measured by extensive next-generation sequencing or by mutation profiles from the point of view of a molecular smoking signature. “Both of these factors are associated with an improved response or can select a group of patients who seem to respond better to immunotherapy.” Mismatch repair genes and microsatellite instability also have been shown to be associated with improved response in other tumor types. “So at least some of these alternatives may well be usable factors for us in the future.”

Optimally, a biomarker would indicate either a high probability of a patient benefiting from a drug or no probability of benefit, he said. Unfortunately, “PD-L1 IHC represents a biological continuum of protein expression from very low levels through moderate to very high levels. So where do we define our positive and negative groups? We know that different drugs and different trials at different times have used different cutoffs defining positive and negative groups of patients.”

This variability creates a biological scenario that is completely different from the addictive oncogene situation, Dr. Kerr explained. “If we use a cutoff, say, of 50 percent to define a group of patients to whom we would give the drug, the patient who has 52 percent as deemed by the pathologist is going to get the drug. How different is that patient from one who is deemed to have a score of, say, 48 percent who is not going to get the drug? Of course we know the answer to this. Biologically, there is really no significant difference between these two patients, but one will be treated and one will not.”

“This is one of the fundamental reasons why PD-L1 does not look as good to our oncologist colleagues when they compare it to the performance of *EGFR* mutations or *ALK* translocation. We have a kind of dose-response relationship to the amount of PD-L1,” which has been consistently shown in all the trials.

Biomarker 'positivity': present, absent, or graduated?



Moreover, PD-L1 is heterogeneous and dynamic. In fact, Dr. Kerr said, "Somewhat paradoxically, if you have a large area of tumor to examine with one of these stains, heterogeneity of expression is actually the norm." Pathologists should be thankful that intensity of expression is not part of this particular biomarker, he added. "It is the proportion of tumor cells that show any staining that is considered the thing we should measure."

Still, "Heterogeneity and dynamism of expression are not friends as far as the pathologist is concerned when measuring this expression. Sampling error is going to occur, that is for sure." These factors might have greater impact on lower thresholds, he said, making the biomarker perhaps appear worse than it really is.

The black marks against PD-L1 are easily summarized, Dr. Kerr demonstrated. "Is the drug targeted in this scenario a singular factor in our target system? Absolutely not. Is the biomarker present or absent? No. Is the biomarker stable and functionally unique? No. Is the biomarker easily measured? Well yes, it is, relatively speaking. But it's not 100 percent predictive."

PD-L1's variability, in fact, makes its potential use with other cancers, such as urothelial cancer or squamous cell carcinoma of the head and neck, unclear. "PD-L1 expression shows predictive value in some cancers but not in others. It's a very curious mixture of signals that we are receiving, and for some oncologists, this is a sign that it's not a very good biomarker."

However, "At the moment PD-L1 IHC is all that we really have to go by. We have to make the best of it for the five drugs now available or expected to come along," he said.

The problem is, "With each drug, we have trials using a different clone, variably different detection systems, and variably different definitions of positivity or scoring mechanisms. This is a huge headache for the pathology world. How are we going to deal with this scenario of four or five drug-assay combinations?"

For one thing, laboratories often have only one staining platform available. It's unlikely that most laboratories will provide the multiple platforms that would be required, he said. "Other big questions concern how different these assays actually are, and the possibility of laboratory-developed tests, which could be built around trial-validated clones or completely different antibody clones."

The outcome of any IHC test is a function not only of the primary antibody that is used but also the detection system and all the chemistry that goes behind it, Dr. Kerr said. "So what chance is there that we might have one test to select a variety of different drugs? Is all PD-L1 IHC the same? The answer is definitely no."

However, he suggested, there is some possibility that one IHC test could be used but scored in multiple different ways to select drugs that are predicated on different definitions of positivity. If that were the case, "pathologists might have to report the PD-L1 IHC according to the number of positive cells indicating a variety of thresholds, and

we might actually have to mention individual drugs in our report—something which, in general, we have not done to date.”

For those contemplating developing a test from scratch, “I think we all have to ask ourselves: How safe is it to deviate from trial-validated practice with our current state of knowledge? I don’t have the answer to that question.”

A study known as the BLUEPRINT project, initiated by the International Association for the Study of Lung Cancer and in which several major cancer groups, diagnostic companies, and pharmaceutical firms were involved, found that three of four assays were remarkably similar when the staining of tumor cells was compared, with a close correlation between 28-8, 22C3, and SP263 results. Similar results have been reported in other studies. However, it is difficult to draw conclusions from these data at the moment, Dr. Kerr said.

Immunotherapy drug availability is another factor. “But as we move these drugs into first-line therapy, I think the dynamic around testing will change, and this will change the way pathologists will have to approach this particular marker.”

In the context of biomarker testing, Dr. Kerr said, after the diagnosis of lung cancer is made, “biomarker testing based on IHC may have to incorporate PD-L1 IHC into the overall testing strategy as part of the panoply of tests we may have to deliver. The biological rationale for this therapeutic approach is that the biomarker test is based on our understanding of differences in antigenicity, evidence of immune response, and evidence of an inhibitory mechanism that may be interrupted by the therapy.”

There is skepticism around immunohistochemistry within the oncology community, Dr. Kerr acknowledged. “We’ve had some rough times in the lung cancer world, and I am often reminded of what happened with HER2 testing in breast cancer. That was not a happy situation in the early days, but I think it’s now very much better.”

For PD-L1 testing, he stressed the value of two things: consistent, quality-assured materials for use in testing, and education and training in how to read the slides. Despite the challenges, he noted, a German study found that after training, pathologists who scored cases in 10-percent brackets had very good concordance. “I cannot emphasize enough the importance of training before you begin.”

For the majority of patients, Dr. Kerr said, it’s obvious whether they are above or below the threshold. But even if the test is carried out exactly in the appropriate way, “the human eye and brain are going to have an issue around the allocation of a case that is very close to the threshold.” With a continuous scale, “there is always going to be a gray zone around the threshold and we have to make the call. And we do our best to make that call using all the techniques and skills and experience we’ve developed.”

PD-L1 IHC is a realistic marker, Dr. Kerr concluded. “But in a very complex environment with multiple drugs and assays, there is no doubt that PD-L1 testing presents us with issues. We have to be practical and realistic in our expectations of this particular biomarker.”

[hr]

Anne Paxton is a writer and attorney in Seattle.