Immunohistochemistry for predictive biomarkers in non-small cell lung cancer

Mari Mino-Kenudson

Department of Pathology, Massachusetts General Hospital & Harvard Medical School, Boston, MA, USA

Correspondence to: Mari Mino-Kenudson. Department of Pathology, Massachusetts General Hospital, 55 Fruit Street, Warren 122, Boston, MA, USA. Email: mminokenudson@partners.org.

Abstract: In the era of targeted therapy, predictive biomarker testing has become increasingly important for non-small cell lung cancer. Of multiple predictive biomarker testing methods, immunohistochemistry (IHC) is widely available and technically less challenging, can provide clinically meaningful results with a rapid turn-around-time and is more cost efficient than molecular platforms. In fact, several IHC assays for predictive biomarkers have already been implemented in routine pathology practice. In this review, we will discuss: (I) the details of anaplastic lymphoma kinase (ALK) and proto-oncogene tyrosine-protein kinase ROS (ROS1) IHC assays including the performance of multiple antibody clones, pros and cons of IHC platforms and various scoring systems to design an optimal algorithm for predictive biomarker testing; (II) issues associated with programmed death-ligand 1 (PD-L1) IHC assays; (III) appropriate pre-analytical tissue handling and selection of optimal tissue samples for predictive biomarker IHC.

Keywords: Anaplastic lymphoma kinase (ALK); proto-oncogene tyrosine-protein kinase ROS (ROS1); programmed death-ligand 1 (PD-L1); immunohistochemistry (IHC); predictive biomarker; non-small cell lung cancer (NSCLC)

Submitted Jun 20, 2017. Accepted for publication Jul 18, 2017.

doi: 10.21037/tlcr.2017.07.06

View this article at: http://dx.doi.org/10.21037/tlcr.2017.07.06

Introduction

In the old paradigm of lung cancer treatment, patients with advanced non-small cell lung cancer (NSCLC), irrespective of histologic subtypes, were typically treated with chemotherapeutic agents, in particular platinum doublets (1). The paradigm of NSCLC treatment started shifting at the beginning of this century, and the discovery of epidermal growth factor receptor (EGFR) mutations in NSCLC paved the way to molecularly targeted therapy and development of predictive biomarkers (2-5). Predictive biomarkers are defined as markers for which the results are essential for therapeutic decision-making, and for treatment with an EGFR tyrosine kinase inhibitor (TKI), the presence of sensitizing EGFR mutations serves as a predictive biomarker. The role of predictive biomarker assays for NSCLC was established in 2011 when the US Food and Drug Administration (FDA) approved both drug and its companion diagnostic test [crizotinib and break-apart fluorescence in situ hybridization (FISH) for anaplastic lymphoma kinase (ALK)] for treatment of patients with advanced ALK-rearranged NSCLC (6,7). Crizotinib, a small molecule TKI of c-MET, ALK and proto-oncogene tyrosine-protein kinase ROS (ROS1), was also approved for treatment of advanced ROS1-rearranged NSCLC in 2016 (8). Thus, break-apart FISH has become the reference standard to detect ALK and ROS1 rearrangements (9,10). However, given the low incidence of ALK (5%) and ROS1 (1-2%) rearrangements in NSCLC (11), expensive FISH assays may not be cost efficient. Thus, immunohistochemistry (IHC) with a sensitive antibody clone targeting ALK or ROS1 protein has been developed as a predictive biomarker assay (12,13).

More recently, the blockade of immune checkpoints to reestablish host antitumor immunity has been investigated extensively in the field of lung cancer, and a few anti programmed cell death protein 1 (PD-1)/programmed
death-ligand 1 (PD-L1) agents have been approved by the US FDA for treatments of advanced NSCLC as the first line or second or more line therapy. In the clinical trials of anti PD-1/PD-L1 agents, PD-L1 IHC assays have been used for predictive biomarker testing, and positive results indicate the presence of an immunomodulatory molecule that can be impacted by the PD-1/PD-L1 blockade (14-20).

In this review, IHC to detect ALK and ROS1 rearrangements and other molecular targets as well as PD-L1 expression will be discussed. Of note, it is important to differentiate the therapeutic decision-making role of IHC for predictive biomarkers from IHC performed for diagnostic purposes, which plays a diagnostically supportive or decisive role and can be critical in distinguishing NSCLC subtypes (21).

**IHC for molecular targets**

**Testing for ALK rearrangements**

ALK rearrangements in lung cancer consist mostly of echinoderm microtubule-associated protein-like 4 (EML4)—ALK translocations (22). First described by a Japanese group led by Dr. Mano in 2007, the EML4-ALK fusion results from a small inversion within the short arm of chromosome 2 leading to expression of a chimeric tyrosin kinase. The chimeric protein has been shown to possess potent oncogenic activity in vitro and in vivo (23). ALK rearrangements have been found in approximately 5% (1–15%) of patients with NSCLC (24).

There are several methods that have been used to detect ALK rearrangements, namely, FISH, IHC, multiplex real time polymerase chain reaction (RT-PCR) and next-generation sequencing (NGS). Of those, multiplex RT-PCR can identify all known ALK rearrangements in a single experiment, and the presence of fusion transcripts as detected by RT-PCR provides direct evidence of chromosomal rearrangements. It requires, however, high-quality RNA, which is difficult to extract from formalin-fixed paraffin-embedded (FFPE) samples. Furthermore, RT-PCR can only detect fusion transcripts with known fusion partners (25). NGS is very efficient to detect rearrangements not only of ALK but also of multiple other genes in a single FFPE sample, although the sensitivity of NGS to detect the gene rearrangements varies among the platforms. Hybrid capture-based NGS can detect most genomic breakpoints, which may be located in introns, while targeted DNA-based NGS methods can detect gene rearrangements only when their breakpoints are adequately covered (22). In order to improve the sensitivity of NGS, anchored multiplex PCR (AMP) has recently been introduced. AMP, a rapid target enrichment method for NGS, is compatible with low nucleic acid input from FFPE specimens, and is effective in detecting gene rearrangements without prior knowledge of the fusion partners (26).

Unfortunately, however, these molecular techniques are not available in many routine pathology laboratories, and their turn-around-time is typically 2–3 weeks that may be too long for patients with rapidly progressive, advanced lung cancer to wait. ALK FISH is currently considered as the universally accepted reference standard, and is approved by the US FDA as a companion diagnostic kit for crizotinib (7). However, the close proximity of EML4 and ALK genes on the short arm of chromosome 2 and other technical or biological conditions may rarely produce equivocal or erroneous results, even when FISH testing is performed with careful preparation and interpretation strictly adherent to the guidelines (22). Furthermore, ALK FISH assays do not seem to be cost efficient in detecting one ALK positive case out of 20 patients with NSCLC. Conversely, IHC is a routine methodology in most pathology laboratories to detect a protein of interest, and appears to be a cost efficient method for reflex testing.

**IHC for ALK**

Initially, several different IHC assays were used for detecting ALK protein expression secondary to ALK rearrangements in the lung. In the published studies, the type or source of antibodies, and antigen retrieval, antibody detection, and amplification techniques varied substantially (22). ALK protein is not expressed in normal lung parenchyma, but is usually expressed secondary to gene rearrangements and the promoter of the fusion partner likely drives the expression. EML4, the most common fusion partner for ALK in lung cancer, encodes a protein that is expressed at a low level in normal lung (27). Of the four most commonly used antibody clones, the clone ALK1 is not sensitive enough to detect the low-level expression of ALK protein secondary to gene rearrangements (Table 1) (12,28,29). Furthermore, increasing the sensitivity with an enhanced detection system will likely result in the greatest amount of background staining with ALK1 (12,45), thus the clone is not preferred as a screening tool. Conversely, a novel monoclonal anti-ALK antibody 1A4 (Origene, Rockville, MD, USA) appears to have high sensitivity compared to the other clones and was considered
Table 1 Performance of various ALK antibody clones in detecting ALK rearrangements in lung cancer*

<table>
<thead>
<tr>
<th>Antibody clone</th>
<th>Study reference</th>
<th>n</th>
<th>Country</th>
<th>Histologic type</th>
<th>Scoring</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Note for positivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK1</td>
<td>Yi 2011 (28)</td>
<td>101</td>
<td>USA</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>90.0</td>
<td>97.8</td>
<td>≥2+</td>
</tr>
<tr>
<td></td>
<td>Walander 2012 (29)</td>
<td>46</td>
<td>USA</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>44.4</td>
<td>67.6</td>
<td>≥1+, adjusted with RT-PCR results</td>
</tr>
<tr>
<td></td>
<td>Mino-Kenudson 2010 (12)</td>
<td>153</td>
<td>USA</td>
<td>Adenocarcinoma</td>
<td>Binary</td>
<td>67.0</td>
<td>97.0</td>
<td>–</td>
</tr>
<tr>
<td>5A4 (Abcam)</td>
<td>Hofman 2012 (30)</td>
<td>20</td>
<td>France</td>
<td>Adenocarcinoma</td>
<td>Binary</td>
<td>50.0</td>
<td>95.7</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>McLeer-Florin 2012 (31)</td>
<td>441</td>
<td>France</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>100.0</td>
<td>98.3</td>
<td>≥1+ in 81 selected specimens</td>
</tr>
<tr>
<td></td>
<td>Lopes 2012 (32)</td>
<td>62</td>
<td>Brazil</td>
<td>NSCLC</td>
<td>Binary (cut-off 10%)</td>
<td>100.0</td>
<td>100.0</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>To 2013 (33)</td>
<td>373</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–3+ based on intensity &amp; extent</td>
<td>100.0</td>
<td>100.0</td>
<td>3+, adjusted with RT-PCR results (only 3+ was seen)</td>
</tr>
<tr>
<td></td>
<td>Cabillic 2014 (34)</td>
<td>1,843</td>
<td>France</td>
<td>NSCLC</td>
<td>0–3+</td>
<td>71.7</td>
<td>99.2</td>
<td>≥1+</td>
</tr>
<tr>
<td></td>
<td>Cabillic 2014 (34)</td>
<td>1,401</td>
<td>France</td>
<td>Predominantly adenocarcinoma</td>
<td>0–3+</td>
<td>67.1</td>
<td>99.4</td>
<td>≥1+</td>
</tr>
<tr>
<td>5A4 (Novocastra)</td>
<td>Paik 2011 (35)</td>
<td>453</td>
<td>Korea</td>
<td>NSCLC (test set)</td>
<td>0–3+</td>
<td>100.0</td>
<td>98.4</td>
<td>≥2+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>187</td>
<td>Korea</td>
<td>Adenocarcinoma (validation set)</td>
<td>–</td>
<td>100.0</td>
<td>98.3</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Kim 2011 (36)</td>
<td>465</td>
<td>Korea</td>
<td>NSCLC</td>
<td>0–3+</td>
<td>100.0</td>
<td>98.1</td>
<td>≥2+ vs. CISH</td>
</tr>
<tr>
<td></td>
<td>Park 2012 (37)</td>
<td>262</td>
<td>Korea</td>
<td>non-squamous NSCLC</td>
<td>0–3+</td>
<td>100.0</td>
<td>98.7</td>
<td>≥1+</td>
</tr>
<tr>
<td></td>
<td>Sholl 2013 (38)</td>
<td>183</td>
<td>USA</td>
<td>N/A</td>
<td>0–2+</td>
<td>92.9</td>
<td>100.0</td>
<td>≥1+, adjusted with clinicopathologic data</td>
</tr>
<tr>
<td>D5F3</td>
<td>Mino-Kenudson 2010 (12)</td>
<td>153</td>
<td>USA</td>
<td>Adenocarcinoma</td>
<td>Binary</td>
<td>100.0</td>
<td>99.0</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Martinez 2013 (39)</td>
<td>79</td>
<td>Spain</td>
<td>NSCLC</td>
<td>Binary (cut-off 10%)</td>
<td>83.3</td>
<td>100.0</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Minca 2013 (40)</td>
<td>249</td>
<td>USA</td>
<td>NSCLC</td>
<td>Binary</td>
<td>100.0</td>
<td>100.0</td>
<td>Adjusted with the 2nd FISH results &amp; clinicopathologic data</td>
</tr>
<tr>
<td></td>
<td>Ying 2013 (25)</td>
<td>196</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>100.0</td>
<td>95.0</td>
<td>≥1+</td>
</tr>
<tr>
<td></td>
<td>Zhou 2014 (41)</td>
<td>368</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>100.0</td>
<td>98.8</td>
<td>≥2+, adjusted with RT-PCR results</td>
</tr>
<tr>
<td></td>
<td>Shan 2014 (42)</td>
<td>286</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–2+ based on intensity &amp; extent</td>
<td>100.0</td>
<td>98.8</td>
<td>≥1+, adjusted with RT-PCR results</td>
</tr>
<tr>
<td>1A4</td>
<td>Wang 2016 (43)</td>
<td>595</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>75.9</td>
<td>99.8</td>
<td>≥1+</td>
</tr>
<tr>
<td></td>
<td>Gruber 2015 (44)</td>
<td>218</td>
<td>Germany</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>100.0</td>
<td>99.0</td>
<td>≥1+, adjusted with RT-PCR results</td>
</tr>
<tr>
<td></td>
<td>Wang 2016 (43)</td>
<td>595</td>
<td>China</td>
<td>Adenocarcinoma</td>
<td>0–3+</td>
<td>100.0</td>
<td>70.3</td>
<td>≥1+</td>
</tr>
</tbody>
</table>

*, compared with ALK fluorescence in situ hybridization (FISH) results. NSCLC, non-small cell lung cancer; N/A, not specified; RT-PCR, reverse transcription polymerase chain reaction; CISH, chromosomal in situ hybridization.
as a promising candidate for ALK rearrangement screening in lung cancer (44); however, it has been shown that the clone 1A4 antibody suffers from lower specificity (70%) (Table 1) (43). Thus, tumors that are positive for ALK with 1A4 IHC require confirmation by another technology such as FISH or NGS. The performance of the other two clones, 5A4 and D5F3, in detecting ALK rearrangements appear to be equally good. As laboratory developed tests, the clone 5A4, in particular that produced by Novostra (currently Leica Biosystems), carries very high sensitivity and specificity (92.9–100% and 98.1–100%, respectively) equivalent to, if not better than, those of the clone D5F3 (Cell Signalling Technology, Billerica, MA, USA) (75.9–100% and 95.0–100%, respectively) (Table 1) (12,25,30–43).

Conklin and colleagues conducted head-to-head comparison of different ALK IHC assays and confirmed that the clones D5F3 and 5A4 (Novostra, Newcastle, United Kingdom) with the ADVANCE system (Dako, Carpenteria, CA, USA) outperformed ALK1 based assays (45). These results are also supported by the recent analysis of pooled data on the diagnostic operating characteristics in 12 studies (3,754 NSCLC specimens) that had evaluated ALK IHC with the 4-tiered scoring systems (46). The sensitivities of D5F3 and 5A4 antibodies were much higher than that of ALK1. With both D5F3 and 5A4 IHC assays, binary IHC scoring with 3+ as the cut-off for positivity, the ALK IHC-positive and negative categories corresponded to ALK FISH-positive and negative cases, respectively. The nearly 100% concordance in these IHC categories supports the use of IHC as a screening method to identify ALK-rearranged NSCLC. However, when tumors exhibit 1+ or 2+ intensity, the results need to be confirmed with ALK FISH or NGS (46). Of note, the reproducibility of ALK IHC results among different laboratories and pathologists is high for any validated protocols (47–50).

Now, an immunoassay [Ventana ALK (D5F3) CDx Assay, Ventana Medical Systems, Tucson, AZ, USA] with the clone D5F3 coupled to the automated immunostaining platform BenchMark XT has been developed and approved by the US FDA as a companion diagnostic kit for crizotinib (51). The assay includes the amplification step with the OptiView Amplification Kit which is intended to reduce or eliminate equivocal results by increasing the signal difference between the specific immunoreaction from the non-specific background staining (that may result in false positive interpretation), thus enabling binary scoring (49). Similarly, with tyramide enhancement that works for both the D5F3 and 5A4 clones, the difference in epitope concentration between negative and strongly positive staining intensities is reduced to the extent that scoring is either negative or positive (22,49,52).

The strong IHC amplification systems used in ALK IHC assays, however, may be associated with various artefacts leading to false-positive results. Positive ALK IHC typically shows strong granular cytoplasmic staining, while weaker cytoplasmic staining may be seen in alveolar macrophages, neural elements, glandular epithelium, extracellular mucin, and areas of necrotic tumor as well as ALK-negative NSCLC (22,49). In addition, some neuroendocrine carcinomas reportedly exhibit positive reactions despite the absence of ALK rearrangements (27,53). Conversely, the signet ring cell morphology that is frequently seen in ALK-rearranged adenocarcinomas could be a source of false-negative staining. A thin membranous positive pattern on ALK IHC may be masked by an intracellular mucin vacuole (54–56). Further, the paranuclear dot-like pattern reportedly associated with the KIF5B-ALK rearrangement may be considered as an artefact, leading to a false-negative result (57). It is very important to understand these possible pitfalls when evaluating ALK IHC, in particular as a standalone test.

Multiple studies have reported discordant results between ALK FISH and IHC, both IHC (D5F3/5A4)+/FISH− and IHC−/FISH+ results, in NSCLC (25,31,33,37,38,40–42,47,50,52,58–68). The discordant results could be attributed to false negative or positive interpretation of FISH results, false negative or positive interpretation of IHC or yet undetermined mechanisms. False-negative interpretation of FISH results could be due to the paucity of tumor cells in the sample, the presence of reactive normal cells interpreted as malignant cells (38), or the physical close proximity of ALK and EML4 genes on the short arm of chromosome 2 (Figure 1) or complex rearrangements involving the ALK gene leading to narrow splits (28). Atypical FISH patterns including solitary green signals with split 5’ centromeric probe could be mistakenly interpreted as positive (38,48). Amplifications of the ALK gene may be associated with ALK protein expression (typically 1+ or 2+ staining) in some cases (69–71). False-positive interpretation of ALK IHC results may be attributed to non-specific high background (27), while suboptimal tissue preservation and fixation with variation of ALK protein expression among specimens could lead to false-negative results (50,59).

Importantly, ALK FISH positivity ranging between 10% and 20% is prone to false negative or positive results, thus it should be interpreted with caution (66). It is worth
mentioning a recent study by Pekar-Zlotin and colleagues that applied NGS as a gold standard for ambiguous cases. They identified six cases with FISH/IHC discrepant results (five IHC+/FISH− and one IHC−/FISH+) in 51 lung adenocarcinomas. NGS confirmed the IHC results in five of the six discrepant cases (except for one with weak IHC+/FISH− results), leading to a sensitivity and specificity of 42.9% and 97.7% for FISH and 100% and 97.7% for IHC compared to an NGS-based approach. Based on the results, they concluded that FISH-based testing may miss a significant number of patients who are eligible for ALK inhibition, and suggested an NGS-based approach in cases with inconclusive IHC staining (64).

ALK IHC assays, already recommended by organizations in Europe, Asia and the United States, are validated and standardized and have been implemented in daily practice as a cost-effective screening tool or a standalone test for detecting ALK rearrangements in NSCLC. When the laboratory conducts an ALK IHC assay as a laboratory developed test (LDT), an optimal algorithm for ALK testing (positive or equivocal results confirmed by FISH or NGS) should be designed. It is also important for laboratories to participate regularly in external quality assessment programs to maintain the reliability of assays.

**Testing for ROS1 rearrangements**

**ROS1** is located on chromosome 6q22 and encodes a receptor tyrosine kinase of the insulin receptor family that shares 77% of amino acid sequences of the ATP-binding site of the tyrosine kinase domain with ALK (72). ROS1 fusions were identified as potential driver events in a cell line (HCC78; SLC34A2-ROS1) and an NSCLC patient sample (CD74-ROS1) in 2007 (73). Subsequently, multiple gene rearrangements involving ROS1 have been reported, but two-thirds of those are distributed over three genes: CD74, EZR, and SLC34A2, each of which has two or more fusion patterns (74). The reported incidence of these fusion proteins in NSCLC is generally low and ranges from 1% to 2% (75). ROS1 may be detected by a variety of techniques, including FISH, RT-PCR, NGS and IHC. Currently, a companion diagnostic kit for crizotinib has not been specified by the US FDA, thus any tests that are validated in individual laboratories could be used to detect ROS1 rearrangements in NSCLC, but the majority use FISH and/or more recently NGS (22). Pros and cons of each assay in detecting rearrangements of ROS1 are similar to those of ALK. Given the presence of multiple fusion partners and easier interpretation of break-apart signals due to interchromosomal rearrangements in the majority of ROS1-rearranged NSCLC, FISH has been used in many pathology laboratories (76). However, FISH may not be cost efficient in detecting one ROS1-rearranged case in 50–100 patients with NSCLC, and ROS1 IHC may serve as an effective screening tool in this context.

**IHC for ROS1**

Most studies on ROS1 IHC have used the D4D6 rabbit monoclonal antibody (Cell Signaling Technology, Danvers, MA, USA) applied at dilutions ranging from 1:50 to 1:250 with various antigen retrieval methods and amplification and detection systems in automated instruments or with
ROS1 overexpression in ROS1-rearranged lung cancers is typically cytoplasmic, but staining patterns vary and may be dependent upon the fusion partner. For instance, a granular cytoplasmic staining with focal or diffuse globular aggregates of protein has been associated with the CD74-ROS1 fusion (77), while membranous staining has been observed in tumors with the EZR-ROS1 fusion (77,78). ROS1-rearranged tumors almost always exhibit diffuse protein expression that is typically homogeneous in a staining pattern but staining can vary in intensify from weak to strong in a tumor (Figure 2A). Detection of ROS1 protein expression in ROS1-rearranged tumors with signet ring cells can be challenging since the cytoplasm may be replaced by non-reactive mucin (77), as described in the ALK-rearranged counterpart (27). ROS1 expression may be seen in tumors without ROS1-rearrangements, most often in a focal or patchy, weak pattern (Figure 2B), but in some cases it can be diffuse and/or strong (79-81). In one study, ROS1 expression was present in 80% of adenocarcinomas with mucinous morphology (invasive mucinous adenocarcinomas) that were negative for ROS1 rearrangements (Figure 2C) (77). Thus, ROS1 expression in this tumor subtype should be interpreted with caution. Similarly, ROS1 expression is occasionally present in non-neoplastic hyperplastic type II pneumocytes and in alveolar macrophages (Figure 2D) (82). In most cases, the expression in these cells is weak-to-moderate (1+ to 2+ in intensity) (22,76), while strong granular cytoplasmic staining of osteoclast-type giant cells has been reported in the setting of a bone metastasis (82).

The performance of ROS1 IHC detecting ROS1 rearrangements is not as good as that of ALK IHC, given the only modest specificities reported in some studies (the reported specificities range from 68% to 100%), while the sensitivity is nearly 100% (Table 2) (78,81-90). Therefore, all IHC positive results need to be confirmed by FISH or another testing method before consideration of treatment with a ROS1 inhibitor. Reported discrepancies between FISH and IHC may also reflect false-negative or false-positive results by FISH (77). Given the easier interpretation of break-apart signals...
# Table 2 Performance of various ROS1 immunohistochemistry assays with clone D4D6 in detecting ROS1 rearrangements in lung cancer*

<table>
<thead>
<tr>
<th>Study (reference)</th>
<th>Country</th>
<th>n</th>
<th>Histologic type</th>
<th>Scoring system</th>
<th>Intensity</th>
<th>No. of positive</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Note for positivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sholl 2013 (82)</td>
<td>USA</td>
<td>218</td>
<td>ADC</td>
<td>Intensity (0–3+; 2–3+ based on extent)</td>
<td>30</td>
<td>18 4 8</td>
<td>100.0</td>
<td>92.0</td>
<td>≥2+</td>
</tr>
<tr>
<td>Mescam-Mancini 2014 (81)</td>
<td>France</td>
<td>121</td>
<td>ADC</td>
<td>Intensity (0–3+)</td>
<td>13</td>
<td>2 11 (2–3+)</td>
<td>100.0</td>
<td>96.9</td>
<td>≥2+</td>
</tr>
<tr>
<td>Yoshida 2014 (77)</td>
<td>Japan</td>
<td>270</td>
<td>269 ADC/1 AdenoSq</td>
<td>H score</td>
<td>95 (any)</td>
<td>46 28 21</td>
<td>94.0</td>
<td>98.0/90.0/87.0</td>
<td>H-score ≥150/75%/≥2+</td>
</tr>
<tr>
<td>Cha 2014 (83)</td>
<td>Korea</td>
<td>330</td>
<td>ADC</td>
<td>H score</td>
<td>56 (any)</td>
<td>N/A – –</td>
<td>100.0</td>
<td>86.4</td>
<td>H-score &gt;100 and ≥75%</td>
</tr>
<tr>
<td>Rogers 2015 (84)</td>
<td>Australia</td>
<td>322</td>
<td>Lung cancer (61% ADC)</td>
<td>Intensity (0–3+) &amp; extent (0–3+)</td>
<td>2</td>
<td>1 0 1</td>
<td>33.3</td>
<td>99.7</td>
<td>≥1+</td>
</tr>
<tr>
<td>Boyle 2015 (78)</td>
<td>USA</td>
<td>33</td>
<td>ADC, molecularly annotated</td>
<td>H score</td>
<td>6 (H score ≥100)</td>
<td>– – –</td>
<td>100.0</td>
<td>100.0</td>
<td>H-score ≥100</td>
</tr>
<tr>
<td>Shan 2015 (85)</td>
<td>China</td>
<td>60</td>
<td>ADC</td>
<td>Intensity (0–3+ in ≥10%)</td>
<td>16</td>
<td>4 8 4</td>
<td>100.0/76.9</td>
<td>93.6</td>
<td>≥1+/≥2+</td>
</tr>
<tr>
<td>Cao 2016 (86)</td>
<td>China</td>
<td>183</td>
<td>ADC</td>
<td>Intensity (0–3+; 2–3+ based on extent)</td>
<td>9</td>
<td>2 3 4</td>
<td>100.0</td>
<td>96.7</td>
<td>≥1+</td>
</tr>
<tr>
<td>Wu 2016 (87)</td>
<td>China</td>
<td>238</td>
<td>NSCLC</td>
<td>H score</td>
<td>42</td>
<td>32 7 3</td>
<td>100.0</td>
<td>100/96.1/68.2</td>
<td>H-score ≥150/≥2+/≥60%</td>
</tr>
<tr>
<td>Kao 2016 (88)</td>
<td>Taiwan</td>
<td>256</td>
<td>ADC</td>
<td>Intensity (0–3+)</td>
<td>57</td>
<td>42 10 5</td>
<td>100.0</td>
<td>100.0</td>
<td>≥3+</td>
</tr>
<tr>
<td>Selinger 2017 (89)</td>
<td>Australia</td>
<td>278</td>
<td>ADC</td>
<td>H score</td>
<td>90</td>
<td>80 9 1</td>
<td>100.0</td>
<td>68.2</td>
<td>any (H score 5–300)</td>
</tr>
</tbody>
</table>

*, compared with ROS1 fluorescence in situ hybridization (FISH) results. †, consisting of retrospective (n=53) and prospective (n=165) cohorts; ‡, consisting of retrospective (n=205) and prospective, EGFR/ALK wild type (n=51) cohorts. ADC, adenocarcinoma; AdenoSq, adenosquamous carcinoma.
with ROSI FISH the contribution of such a factor should be limited, however.

Based on our experience and others, diffuse and homogeneous protein expression and/or high H-score may be used in selecting cases for the confirmatory testing. For instance, some studies have suggested that application of H-score cutoffs of 100 or 150 can maximize the sensitivity and specificity of ROSI IHC (77,78). Thus, experienced pathologists may be able to distinguish tumors harboring ROSI rearrangements from those with false-positive IHC results based on the distinctive patterns of ROSI expression. Although individual laboratories may opt to perform confirmatory testing on only those samples with diffuse and homogeneous expression patterns, a confirmatory analysis on all samples examined with IHC is recommended to evaluate the performance of ROSI IHC at least at the beginning of the IHC implementation. In France, a national pathology expert panel recommends a screening with ROSI IHC and confirmatory FISH for IHC positive or equivalent cases to detect ROSI-rearranged NSCLC (81).

**PD-1/PD-L1 blockade in NSCLC**

High profile clinical trials have demonstrated impressive anti-tumor activity of anti PD-1 and PD-L1 agents in NSCLC (19,91,92), and significant improvements in overall survival (OS) of previously treated, advanced NSCLC patients compared to single-agent docetaxel (16,18,20,93). Subsequently, the US FDA approved nivolumab, pembrolizumab and atezolizumab for NSCLC patients with disease progression on or after platinum-based chemotherapy (94). Clinical trials with two other agents, durvalumab and avelumab, have also shown promising results (14,15). Furthermore, pembrolizumab was approved by the FDA as a 1st line treatment for advanced NSCLC patients in October, 2016 (94), based on the results of the clinical trial, KEYNOTE-024, that showed significantly improved objective response rate, progression free survival (PFS), and OS when advanced NSCLC patients, whose tumors harbored PD-L1 expression by IHC in 50% or greater of the tumor cells, were treated with pembrolizumab compared to platinum-based chemotherapy in the 1st line setting (95). In these clinical trials, PD-L1 expression in the tumor by the specific IHC assay served as a predictive biomarker. Importantly, five different PD-L1 IHC assays were developed for the five PD-1/PD-L1 agents (Table 3) (96,97), and some of the assays have been approved by FDA.
as either companion (a requirement for drug eligibility) or complimentary (only for guidance) diagnostic kits along with the corresponding anti PD-1/PD-L1 agents.

Now, anti PD-1/PD-L1 agents are incorporated in the paradigm of treatment for advanced NSCLC patients. Given the availability of pembrolizumab as a first line therapy, the National Comprehensive Cancer Network (NCCN) guidelines recommend that all advanced NSCLC samples be tested with a PD-L1 IHC assay in a reflex manner (98).

**PD-L1 IHC testing**

As briefly mentioned, five different PD-L1 IHC assays (the Dako 28-8, Dako 22C3, Ventana SP142, Ventana SP263 and Dako 73-10) have been developed and validated as a predictive biomarker in the clinical trials (Table 3) (96,97). Pathology laboratories need at least one affordable, validated test, but it is not practical for them to conduct several specific assays for one protein from a financial and regulatory perspective, and given the limited availability of tumor tissue for testing and the number of tissue-based diagnostic tests required in the management of an advanced NSCLC patient. However, selecting one assay among several available is challenging. While each of the five IHC assays recognizes PD-L1 protein, each antibody clone appears to be specific for a different epitope of the PD-L1 protein and may not have the same binding affinity for its epitope. In addition, different detection systems, with or without amplification, are used in different assays, thus the performance of these five assays may be different. Scoring systems to determine “positive” results are also quite different between the five assays, since these systems were determined based on a predictive value and clinical data obtained during the therapeutic/diagnostic test co-development process of individual anti PD-1/PD-L1 agents. For instance, complete circumferential or partial linear membranous staining of tumor cells irrespective of intensity is considered positive for the DAKO 22C3, 28-8 and 73-10 assays (Figure 3A,B), while any membranous and/or cytoplasmic expression of tumor cells is considered positive for the Ventana SP263 assay. Ventana SP142 assay is unique...
in that not only tumor cell expression, but also immune cell expression, are taken into account. Furthermore, each assay has a specific % of positive tumor cells as a cut-off, and the percent may be different depending on the 1st line vs. 2nd or more line of treatment. Importantly, the outcomes of patients, when using these drugs in a cohort selected for “positive” PD-L1 expression, have only been evaluated in trials using the specific drug-assay combinations (Table 3).

In order to see whether we can use only one of these assays or any other laboratory developed tests (LDTs) to select patients for all available anti PD-1/PD-L1 agents, several studies, although only from the technical/analytical perspective, have compared the performance of clinical trial PD-L1 IHC assays (99-101). The International Association for the Study of Lung Cancer (IASLC) and the American Association for Cancer Research (AACR), in collaboration with pharmaceutical companies and diagnostics vendors, have evaluated the technical similarities and differences of the Dako 28-8, Dako 22C3, Ventana SP263, and Ventana SP142 assays (101). The initial phase I part of this study was to test the feasibility on a small cohort of 38 NSCLC resections (not treated with anti PD-1/PD-L1 agents), each stained using all four assays and scored by three trained pathologists. Of those, the Dako 22C3, Dako 28-8, and Ventana SP263 assays demonstrated the similar membranous staining on tumor cells, while the Ventana SP142 assay consistently revealed smaller numbers of positive tumor cells. IASLC is now conducting a larger phase II study consisting of different sample types (resection, biopsy and cell block) stained using 5 clinical trial assays (including the Dako 73-10 assay) evaluated by more than 20 pulmonary pathologists. The German ring study assessed interobserver concordance and PD-L1 IHC staining patterns in 15 NSCLC resection specimens using the four PD-L1 IHC assays (the Dako 28-8, Dako 22C3, Ventana SP263, and Ventana SP142 assays), and showed that the tumor cell score could be reproducible, with no differences in interobserver concordance among the tested assays (100). The scoring of immune cells, however, yielded low concordance rates indicating that immune cell scoring might require specific standardization. In addition, they found that staining patterns/intensities might be different among the four assays. While the 28-8 and 22C3 assay stained similar proportions of tumor cells in the majority of cases, the SP142 assay stained fewer tumor cells and SP263 stained more tumor cells compared to the 28-8 and 22C3 assays in some cases. Another recent study, in which 493 NSCLC tissue samples were stained using three assays (the Dako 28-8, Dako 22C3 and Ventana SP263 assays) and scored by a single trained pathologist, has reported the similar patterns of tumor membranous staining with high (>90%) overall percentage agreement between the assays at multiple expression cut-offs (including 1%, 10%, 25% and 50%) (99).

The results of these studies raise the possibility that the Dako 22C3, Dako 22-8 and Ventana SP263 assays can be used interchangeably to identify patients most likely to respond to anti PD-1/PD-L1 agents (nivolumab, pembrolizumab and durvalumab) provided that the appropriate scoring system is used for the corresponding agent (99). Unfortunately, however, the Ventana SP142 assay that consistently shows fewer numbers of positive tumor cells and includes the immune cell component for scoring does not appear to be interchangeable with the other assays. What about the performance of LDTs that could be developed with either clinical trial antibody clones or non-clinical trial clones? A recent study by Neuman and colleagues has reported successful implementation of 22C3 IHC on the Ventana Benchmark XT platform with two of the Ventana’s detection systems after a rigorous optimization process (102). Among multiple non-clinical trial PD-L1 antibody clones, the clone E1L3N (CST) has been optimized with various IHC platforms and detection systems, and have been used in multiple clinical studies (100,103-113). Importantly, a prospective, multi-institutional study sponsored by NCCN and Bristol-Myers Squibb has shown the possible utility of E1L3N IHC on the Leica Bond platform (114). The study, in which 13 pathologists scored 90 surgically resected NSCLCs stained using the Dako 22-8, Dako 22C3, Ventana SP142 and E1L3N/Leica assays have shown analytical equivalency of the Dako 28-8 and E1L3N/Leica assays using the average pathologist scores across all 90 cases. While the Dako 22C3 assays revealed significantly less expression than the other two antibodies (only when averaging the readings of 13 pathologists), there was no difference in sensitivity and specificity equivalents between the Dako 22-8, Dako 22C3 and E1L3N/Leica assays using a “real world” assessment (agreement by individual pathologists). The Ventana SP 142, however, again revealed significantly less expression by large amount with every method of assessment (114). The results of these studies have brought optimism that harmonization between assays, including LDTs, may be possible. Of somber note, the recent French study comparing performance of clinical trial assays and various combination of LDTs has reported only half of the
evaluated LDTs demonstrating sufficient concordance with the reference assays for tumor cell scores (115). The study consisted of 41 resected NSCLCs stained with clones 28-8, 22C3, SP263, SP142 and E1L3N in 7 centers (3 with Dako Link 48, 2 with Ventana Benchmark Ultra and 2 with Leica Bond III; 8 clinical trial assays and 27 LDTs) scored by 7 trained pathologists, and found that only 14 of the 27 LDTs achieved >0.75 weighted kappa coefficient compared to the reference assay (115). Thus, not all LDTs will be interchangeable with the clinical trial assays.

As for interobserver concordance on PD-L1 tumor cells scoring, in the German Ring study with highly selected cases, the overall percent agreement at the 1% threshold ranged for all 4 assays from 90.4% to 97.2%, and at the 50%, from 91.5% to 94.8% (100). An Australian study by Cooper and colleagues using the Dako PD-L1 22C3 assay on a highly selected set of cases showed the overall percent agreement at the 1% threshold of 84.2%, and at the 50% threshold of 81.9% (116). Rehman and colleagues examined the reproducibility of 5 pathologists on a selected set of cases stained with the Ventana SP142 assay and showed an intraclass correlation coefficient of 94% agreement among the pathologists for the assessment of PD-L1 in tumor cells, but only 27% agreement on stromal/immune cell PD-L1 expression. The subjective interpretation of the guidelines for scoring stromal cells may have contributed to the fair agreement on PD-L1 expression on stromal/immune cells (117).

**IHC for other predictive biomarkers**

IHC may be used to detect other predictive biomarkers in NSCLC. Of those, EGFR mutation specific antibodies recognize the protein conformation change due to the mutation, but do not bind to the wild type EGFR protein. There are two types of EGFR mutation specific antibodies—the one specific for a 15 bp deletion in exon 19 and the other for a L858R point mutation in exon 21. Currently, these are not recommended for predictive testing, since the sensitivity for detecting the corresponding mutations is modest, while the specificity is high (118). However, EGFR mutation specific IHC may be useful when available tissue samples are insufficient for molecular assays due to scarcity of tumor cells and/or fixation with decalcification or heavy metal solutions.

In Europe, EMA approved monoclonal anti-EGFR drug, necitumumab, for patients with advanced-stage squamous NSCLC expressing the wild type EGFR protein by IHC (119). The US FDA also approved necitumumab for use in combination with cisplatin/gemcitabine chemotherapy for the first line treatment of metastatic squamous NSCLC, but EGFR protein expression by IHC is not a requirement for the treatment (120).

A small fraction (2–4%) of NSCLC harbors BRAF mutations. A BRAF V600E-mutation specific antibody has been developed and proven useful in detecting BRAF V600E mutations in colon cancer (121), but hardly detects any of the proteins encoded by non-V600E mutations (122,123). In NSCLC 40–50% of the BRAF mutations are non-V600E (124), thus BRAF V600E-mutation specific IHC does not appear to be useful.

NSCLC with MET amplification or exon 14 skipping mutation could be treated with a MET inhibitor, crizotinib (or others) in clinical trials. High MET protein expression by IHC appears to be associated with MET FISH positivity and amplification, but there is a significant overlap between FISH/amplification positive and negative cases. In addition, MET exon 14 skipping mutations are proportionally much rare (125), thus, MET IHC does not seem to be an efficient method for the detection of MET alterations amenable for MET inhibition.

**Pre-analytical variables for IHC**

The stability of proteins can be affected by multiple pre-analytic variables, including ischemia time, fixative type, fixation time, archive conditions, and age of archived material (126). Thus, the successful implementation of IHC-based assays in general depends on pre-analytic tissue handling (127) as well as the antigen retrieval and detection systems. An ischemia time from excision to the initiation of fixation should be short (as short as possible), and biopsies should immediately be immersed in fixative for 6–48 h. Formalin (neutral buffered formalin) is historically the preferred and most common fixative used in the practice of histopathology (128). Consequently, the majority of pathology laboratories typically perform the initial validation of IHC protocols on FFPE tissue. Decalcifying solutions used for bony specimens vary in their effects on retention and integrity of nucleic acids and proteins. Thus, results of IHC on decalcified specimens are unpredictable because of wide variations in specimen types and sizes, fixation time, and the particular solution(s) used (129). Similarly, alcohol fixation used for cytology specimens including alcohol-fixed cell blocks decreases IHC accuracy by causing loss or decrease of
immunogenicity when IHC protocols optimized with FFPE tissue samples are used (126).

In order to improve the accuracy of IHC performed on samples fixed in these fixatives, in the US, the guidelines from the College of American Pathologists (CAP) Pathology and Laboratory Quality Center recommend that pathology laboratories test a sufficient number of such cases to ensure that assays consistently achieve expected results (129). It is of particular importance since at least 30–40% of advanced NSCLCs are diagnosed by cytology alone.

Rigorous validation and protocol optimization should be performed in each laboratory that performs IHC on cytology specimens (e.g., alcohol-fixed cell blocks, air-dried smears, formalin-post fixed specimens) (126). Irrespective of specimen types, the CAP guideline recommends examination of at least 20 positive and 20 negative samples for initial analytic validation of predictive marker assays, while a minimum of 10 positive and 10 negative tissues are sufficient for non-predictive marker assays (129).

**Special note on pre-analytical variables for PD-L1 IHC assays**

The PD-L1 IHC assays have not been validated for decalcified tissue (130,131), thus PD-L1 IHC on decalcified tissues should be avoided when another tissue sample is available. Specimen age for PD-L1 testing should be less than 3 years, since antigenicity may drop significantly in those older than 3 years (132). In addition, in our experience, the antigen retrieval conditions (citrate buffer pH6 vs. citrate buffer pH8 or EDTA pH9) have been shown to significantly affect intensity and rate of positivity for the clone E1L3N (unpublished observation). These differences may alter the outcome of the test, leading to alternative PD-L1 scoring around a given cut-off threshold.

Importantly, the use of cytology samples for PD-L1 IHC is currently not recommended, due to the lack of rigorous validation for this purpose (131). The study by Rebelatto and colleagues has also shown that 95% alcohol, AFA and PRFER are unacceptable fixatives for IHC with the SP263 clone (133). Furthermore, an evaluation of PD-L1 positive immune cells with the Ventana SP142 assay will likely be more challenging in cytology specimens, since the lack of tissue architecture precludes distinction of the relevant immune cells within the tumor area from immune cells outside of the tumor boundaries that are considered irrelevant for PD-L1 scoring. Pre-existing lymphocytes in a fine-needle aspirate of a lymph node also precludes an optimal immune cell scoring (96).

When a cytology specimen is the only available sample, however, a cell block may be used for PD-L1 testing as long as it is FFPE and contains a sufficient number of tumor cells (minimum 50–100) (Figure 3C,D) (134-136). Several studies have reported high concordance on PD-L1 expression between cell blocks and matched histological specimens and/or comparable PD-L1 expression among cell blocks, small biopsies and resections in a prospective cohort (134-136). The results of these studies suggest that cytological materials are as good as histological materials for the tumor cell analysis, but large-scale validation studies are warranted to establish PD-L1 IHC testing on cytology specimens.

**Conclusions**

Given the increasing number of targeted therapies available for advanced NSCLC patients in clinic, predictive biomarker testing has never been more important. IHC is a widely available and technically less challenging method to perform and interpret, can provide clinically meaningful results very quickly and is more cost efficient than molecular assays. Thus, the use of IHC to detect predictive biomarkers, in particular gene rearrangements, in NSCLC has been well established. It is important, however, to understand the performance of multiple antibody clones, pros and cons of IHC platforms and various scoring systems to design an optimal algorithm for predictive biomarker testing. In addition, given the recent FDA approval of an anti PD-1 agent as a first line therapy for advanced NSCLC patients, PD-L1 IHC testing has become routine in pathology laboratories in the US and some other countries. The one drug-one predictive biomarker assay concept, however, has brought unique challenges to pathology and oncology communities, since operating five different PD-L1 IHC assays to support the use of five drugs is extremely difficult in any laboratory from both the practical/regulatory and financial perspective. With this background, several recent studies evaluated and compared the analytical/technical performance of clinical trial assays and/or LDTs leading to the optimism that harmonization of PD-L1 IHC assays may be feasible, if rigorous optimization and validation are performed for LDTs that are used to identify patients who will likely respond to anti PD-1/PD-L1 agents. Finally, appropriate pre-analytical tissue handling and selection of optimal tissue samples are the keys to successful IHC, not only for predictive biomarkers, but also for protein expression in general. Of particular note, the
use of cytology smear and cell blocks for PD-L1 IHC is not currently recommended, but cell blocks that are made following appropriate protocols may be used with caution.

Acknowledgements

The author thanks Mr. Cris Kenudson for editorial support. Funding: This work was supported by a Stand Up To Cancer-American Cancer Society Dream Team Translation Research Grant.

Footnote

Conflicts of Interest: M Mino-Kenudson has served as a consultant for Merrimack Pharmaceuticals, H3 Biomedicine, and Advanced Cell Diagnostics and as an advisory board member for Roche.

References

6. Available online: https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/202570s000lbl.pdf
7. Available online: https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm
8. Available online: https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm490391.htm


45. Conklin CM, Craddock KJ, Have C, et al. Immunohistochemistry is a reliable screening tool for


51. Available online: https://www.accessdata.fda.gov/cdrh_docs/pdf14/p140025c.pdf


89. Selinger CI, Li BT, Pavlakis N, et al. Screening for ROS1 gene rearrangements in non-small-cell lung cancers using immunohistochemistry with FISH confirmation is an effective method to identify this rare target. Histopathology 2017;70:402-11.


94. Available online: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm279174.htm


130. Available online: http://www.accessdata.fda.gov/cdrh_docs/pdf15/P150013c.pdf