

Scoring gastric, GEJ cancers for PD-L1 expression

Anne Paxton

February 2018—To some ears, perhaps, the scientific method connotes a process that is standardized and unimaginative. But inventions like Velcro, vulcanization, and the microwave—all stemming from accidental discoveries—testify to the role of luck and leaps of intuition in formulating and modifying a hypothesis.

When pathologists and scientists at Agilent Technologies, maker of the companion diagnostic PD-L1 IHC 22C3 pharmDx, sought to develop a new scoring methodology, in collaboration with Merck, for gastric cancer patients' tumor specimens a few years ago, luck and intuition turned out to be handy. As Agilent's chief pathologist for companion diagnostics, Debra Hanks, MD, puts it, "Skill, science, and a smidgeon of serendipity" helped her team zero in on the best way to evaluate PD-L1 expression in gastric cancer patients.

The result, called the combined positive score (CPS), is part of a companion diagnostics package approved Sept. 22, 2017 by the FDA. The CPS promises to identify more precisely gastric and gastroesophageal junction adenocarcinoma patients who are likely to respond to the drug pembrolizumab (Merck's Keytruda).

In an Oct. 17 webinar hosted by CAP TODAY with an educational grant from Agilent (available at captodayonline.com), Dr. Hanks describes how pathologists can adopt and perfect their use of the CPS in their own laboratories when analyzing gastric cancer patients' biopsies, with help from Agilent's interpretation manual and Web-based training materials for the diagnostic.

The basic mechanism of PD-L1 is well understood: Tumor cells can use the PD-L1 immune cascade to turn off cytotoxic T-cells so the tumor can continue to grow; the antibody pembrolizumab blocks the PD-L1 so the cytotoxic T-cells can continue to stay active and attack the tumor. But "as we learn more and more about immunotherapy, what we're finding is that different cancer types can express PD-L1 differently," explains Dr. Hanks, who spoke with CAP TODAY.

Early translational work at Merck Research Laboratories, which studied four tumor types, including gastric cancer, revealed that tumor cell PD-L1 expression was lacking in the majority of those responding to Keytruda. This suggested that the tumor proportion score (TPS), which worked so well in non-small cell lung cancer, would not be a useful biomarker in many other cancers. Interestingly, most of these tumors contained PD-L1-expressing immune cell infiltrates. Scoring immune cell infiltrates was notoriously difficult.

Gastric cancer alone is a disease with a dismal prognosis: five-year survival of 30 percent in the U.S. and five percent worldwide, so the stakes of finding the right scoring methodology for gastric cancer, and other tumor types, were high.

The interpretation criteria for the new CPS methodology are part of the full-solution diagnostic package the FDA approved, explains webinar co-presenter Annika Eklund, PhD, global product manager for Agilent's companion diagnostics program. "The FDA approved all reagents in the kit, including the 22C3 primary antibody, licensed from Merck; the EnVision Flex visualization system; the control cell lines; the staining procedure on Autostainer Link 48 (software); the instrument Autostainer Link 48 (hardware); and finally the interpretation criteria used by the pathologists in the clinical trial." Pathologists will find the product insert, Agilent's interpretation manual, and E-Learning modules to be important sources of information and guidance, Dr. Eklund says. Based on the clinical trial outcomes, the FDA granted accelerated approval in heavily treated PD-L1-positive gastric cancer patients. Through use of the diagnostic PD-L1 IHC 22C3 pharmDx, spurred by the FDA approval, patients with gastric or gastroesophageal junction cancer are already being selected for treatment with pembrolizumab.

"When we collected data for the preliminary clinical trial, called the Keynote-012, Agilent and Merck found that the tumor proportion score . . . did not work well to identify gastric patients," Dr. Hanks says. The tumor proportion

score algorithm identified only two of the 11 gastric patients who responded to Keytruda, so a better method was needed. “We were asked to look at slides from patients who were responders and nonresponders to see if there are any particular staining patterns or expression patterns that would identify responders in gastric cancer.”



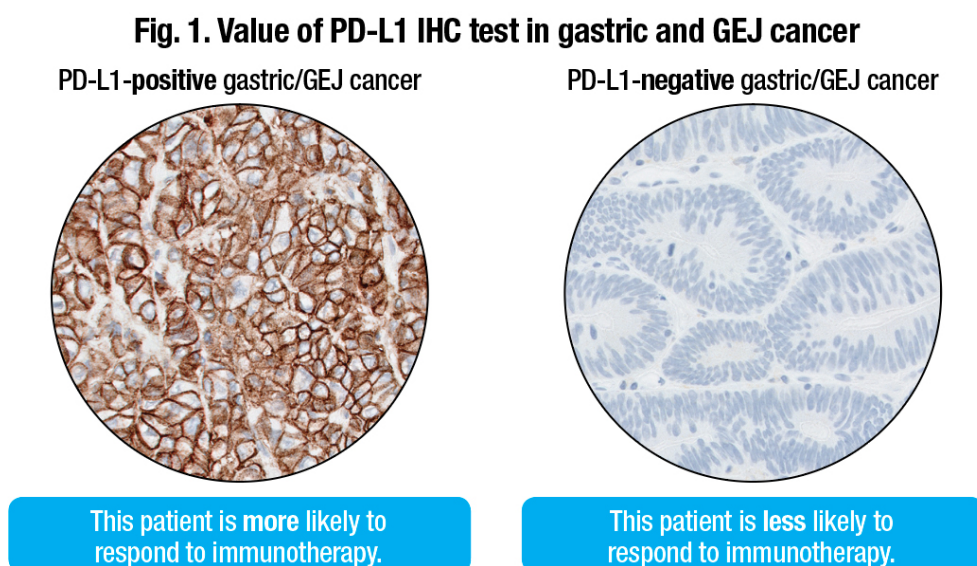
Dr. Hanks

The team recorded different staining parameters observed on the slides, breaking down where staining occurred and the amount and intensity, and tried different ways of scoring to develop the new system before arriving at the combined positive score. “Sometimes the process involved trial and error, but it was mainly putting together different formulas and mechanisms by which you could come up with a score, and then seeing if that identified the responders,” Dr. Hanks says. And the scientists and pathologists on the team hit on the key: They discovered that by counting the tumor cells, lymphocytes, and macrophages relative to the viable tumor cells present, they were able to identify nine out of the 11 responders. The combined positive score worked.

To calculate a CPS, the pathologist must score the number of PD-L1-positive cells (tumor cells, lymphocytes, and macrophages), divide that total by the number of viable tumor cells, and multiply by 100. For gastric or gastroesophageal junction adenocarcinoma, a CPS score ≥ 1 identifies responders who are eligible for treatment with pembrolizumab.

To test the value of this experimental algorithm, another clinical study, a phase two clinical trial called Keynote-059, examined the response of the patients identified by CPS to treatment with Keytruda as a third-line monotherapy. To qualify for Keynote-059, the patients had to have shown disease progression on at least two prior treatment regimens and had to have a life expectancy of at least three months, Dr. Hanks says.

For the 200-plus patients in the Keynote-059 study, about 58 percent had PD-L1 expression (CPS ≥ 1). For the 143 patients who had PD-L1 expression, the overall response rate was 13.3 percent; 1.4 percent had a complete response and 11.9 percent had a partial response (**Fig. 1**).



An important point for pathologists who will be using CPS is the difference between results of testing archival tissue or newly obtained tissue—that is, tissue obtained 42 days before the patient’s first dose of Keytruda. About 49 percent of the cohort studied was positive for PD-L1 if it was archival, but the prevalence rose to 73 percent for newly obtained tissue. Therefore, Dr. Hanks says, “For the drug and our assay label, if you have archival tissue and it tests negative, we recommend that you then obtain and test fresh tissue, to improve the probability of detecting PD-L1.”

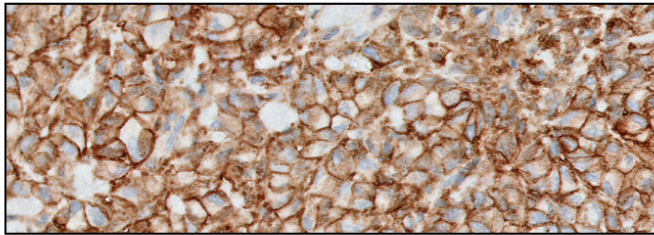
Under CAP requirements, pathologists should save blocks and slides for 10 years, she notes, but with the number of biomarkers being identified for companion diagnostics, archival tissue is often an opportunity for oncologists to call the pathologist and order a new test on the patient. “In the case of NSCLC,” Dr. Hanks says, “we have found that blocks that are five years or older could result in a loss of PD-L1 immunoreactivity.”

Agilent’s interpretation manual and four E-Learning modules online will aid pathologists, Dr. Hanks says, in evaluating specimen adequacy and PD-L1 staining results for gastric cancer, calculating the CPS, reporting results, and testing their expertise by scoring a variety of gastric cases, from simple to complex.

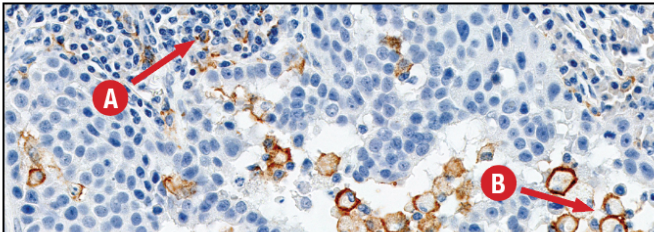
Specimen adequacy, naturally, is a key concern for pathologists. “It’s part of our training in pathology to base a diagnosis on the best tissue,” Dr. Hanks says, “to make sure the tissue is adequate and that it’s properly fixed. It doesn’t happen that frequently that we have to deal with preanalytical issues, however, and PD-L1 produces a very robust IHC stain.”

What cells to include in the numerator of the score, and what cells to exclude, are pivotal parts of evaluating the stain, she says, recommending that pathologists keep a copy of the inclusion and exclusion criteria for the algorithm next to their microscope.

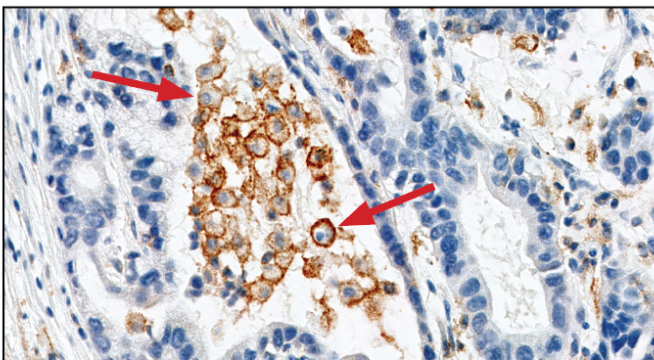
Fig. 2. Tumor and tumor-associated immune cells



PD-L1 primary antibody exhibiting linear membrane staining of tumor cells (20× magnification).



PD-L1 primary antibody exhibiting linear membrane staining of tumor-associated mononuclear inflammatory cells (A: lymphocytes, B: macrophages) (20× magnification). Note: Tumor-associated MICs are present within the tumor nests and/or adjacent supporting stroma and are directly associated with the response against the tumor.



PD-L1 primary antibody exhibiting linear membrane staining of tumor-associated macrophages (arrows) (20× magnification).

Key point: Tumor-associated lymphocytes and macrophages with convincing membrane and/or cytoplasmic staining should be included in the CPS numerator.

She refers to the images in Fig. 2 to illustrate how the criteria should be employed in calculating the numerator. “The upper image shows positive tumor cells in a beautiful staining pattern that ranges from two-plus to three-plus intensity. These are scored in the exact same way as in our algorithm for non-small cell lung cancer. It’s cell membrane staining. We only score tumor cells in the numerator as positive by membrane staining at any intensity, partial or complete, and it has to be convincing. We do not score tumor cells if they only have cytoplasmic staining; they are considered negative.” A 20× objective is used to decide if weak staining is true membrane staining. “Tumor cells with a small arc of membrane staining are scored as positive,” Dr. Hanks explains.

The other images in Fig. 2 show a typical immune cell staining pattern for the PD-L1 22C3 kit. “You can count any of the immune cells, whether they be macrophages or lymphocytes, and include them in the numerator. But the key point is that these tumor-associated lymphocytes and macrophages have to have convincing membrane and/or cytoplasmic staining at any intensity to be interpreted as positive.” (**Fig. 3**).

Sometimes the results can be counterintuitive. “If you only had an H&E, you might review a particular slide and say, ‘This is a poorly differentiated adenocarcinoma of the stomach; this patient has a horrible prognosis.’ But the high amount of PD-L1 staining on a slide of the tumor cells could show the probability that the patient potentially will respond to Keytruda and have an improved prognosis.”

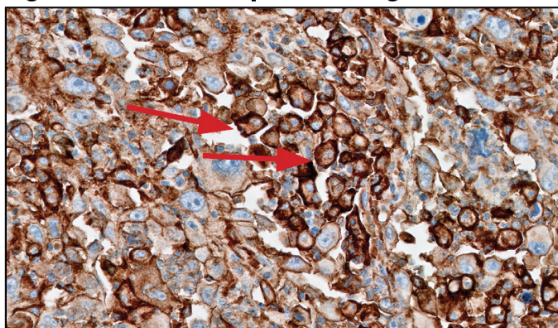
Evaluating the other part of the numerator, the tumor-associated immune cells, can also be difficult. “It’s known from published studies, such as the Blueprint Study”—jointly sponsored by the FDA, American Society of Clinical Oncology, and American Association for Cancer Research to build an evidence base for PD-L1—“that pathologists can have some challenges in scoring immune cells accurately, and so we have worked very diligently with our

team of pathologists and scientists to dissect this out and define how to score immune cells. And we've included a lot of those caveats, pitfalls, and tips in our literature," Dr. Hanks says. **(Fig. 4).**

Deciding what is tumor associated or not is challenging. "We count PD-L1 immune cells associated with the tumor. And when we have a question, we put the tumor in the middle of a 20x field, and any positive immune cells, or what we call mononuclear inflammatory cells [MICs], would be counted. In the same way, if you have a nest of tumor cells and you put it in the middle, any positive membrane and/or cytoplasmic staining at any intensity of the lymphocytes or macrophages are all included in the score." **(Fig. 5).**

It's quite common to see a nest of gastric cancer tumor cells surrounded by PD-L1-positive lymphocytes and macrophages. "We see this staining pattern with PD-L1 quite frequently," she notes. All of those would count. Lymphoid aggregates are another staining pattern that she refers to as the "mother lode" of positive numerator in the lymphocytes. "These are clusters of lymphocytes that can be found within lots of tumors, and a pathologist can identify their morphology very easily at low power if they are really packed together in a small region."

Fig. 3. Partial and complete staining of tumor cells



PD-L1 primary antibody exhibiting linear membrane staining distinct from cytoplasmic staining (arrows) (20x magnification).

However, Dr. Hanks emphasizes, it's important to keep in mind that when the first impression is that the CPS is less than one, "it's not enough to determine if your specimen is indeed CPS zero or PD-L1 expression-negative from a low power scan. You have to review the entire slide at 20x to make sure you're not missing pockets of weakly staining tumor cells or lymphocytes or macrophages."

The CPS exclusion criteria specifically rule out any PD-L1-positive immune cells associated with adenoma, dysplasia, carcinoma, or with ulcers, chronic gastritis, or other processes not associated with the tumor, Dr. Hanks says. "We look at the H&E stain slide of patients' tumors very carefully, and pathologists are able to identify where the tumor cells are versus where adenoma or carcinoma in situ are located."

"We do not count immune cells that are associated with normal structures, and we do not include neutrophils, eosinophils, or plasma cells. In gastric cancer, you can find small clusters of plasma cells and they tend to stain with a very weak cytoplasmic blush. They would be excluded from scoring."

Fig. 4. CPS: Evaluating the numerator. Identifying PD-L1-positive tumor-associated immune cells (lymphocytes and macrophages)

of PD-L1-positive cells (tumor cells, lymphocytes, macrophages)

CPS =

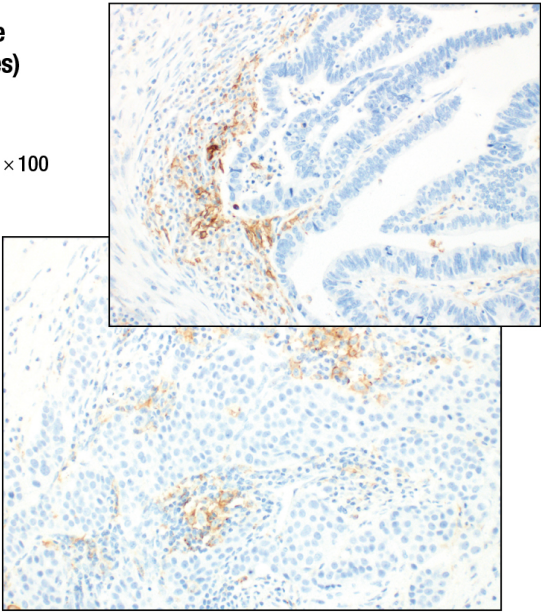
Total # of tumor cells

× 100

Element	Included in scoring
Immune cells	Membrane and/or cytoplasmic* staining (at any intensity) of mononuclear inflammatory cells (MICs) within tumor nests and adjacent supporting stroma:
	■ Lymphocytes (including lymphocyte aggregates)
	■ Macrophages
	Only MICs directly associated with the response to the tumor are scored.

*In MICs membrane and cytoplasmic staining are often indistinguishable due to high nuclear to cytoplasmic ratio. Therefore, membrane and/or cytoplasmic staining of MICs are included in the score.

Immune cells may have smaller nuclei than tumor cells: A larger number of MICs can occupy the same area as a smaller number of tumor cells.

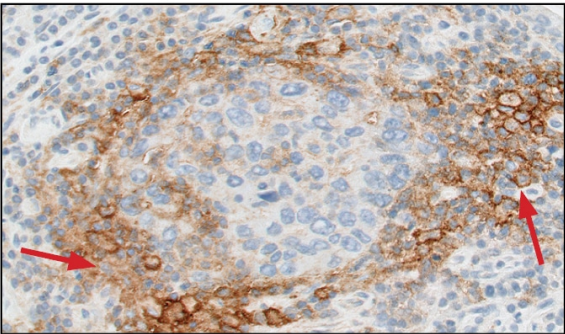


Ganglion cells reflect another ailment that is excluded . These can be found deep in the smooth muscle wall of a gastric resection specimen and

can have a blush stain that is similar to cytoplasmic or membrane staining, but should not be counted. However, “We have also found that a low-power streaming of staining could be interpreted as only fibroblast staining, but when you get to 20× there are some immune cells mixed within fibroblasts.” Also excluded are stromal cells including fibroblasts and, naturally, necrotic cells or cellular debris. Sample images in Agilent literature on the PD-L1 assay, including the interpretation manual and E-Learning modules, explain this aspect of the CPS in greater detail.

Evaluation of the CPS denominator involves a process pathologists are familiar with from scoring estrogen and progesterone receptors, Dr. Hanks says. With CPS, the task is somewhat more quantitative. “All the viable tumor cells on your slide count, so tumor cells with and without staining should be carefully examined to precisely assess the denominator.” For example, with a tumor cell that’s about 20 μm in diameter, there are about 2,500 cells filling a typical 20× field.

Fig. 5. Staining of tumor-associated MICs



PD-L1 primary antibody exhibiting linear membrane and/or cytoplasmic staining of tumor-associated mononuclear inflammatory cells (arrows) (20× magnification).

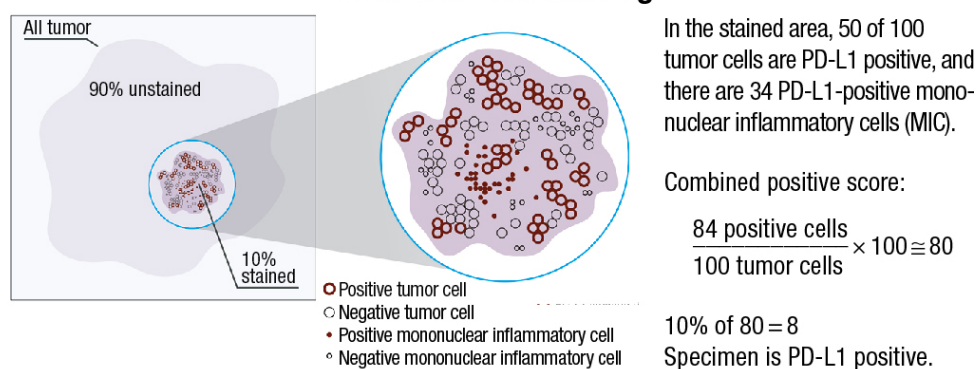
Diffuse gastric carcinoma creates other difficulties for pathologists evaluating the denominator, Dr. Hanks notes. “Those pathologists who frequently look at resections know there can be a dense infiltrate right underneath the

mucosa, and then the cells invade through the smooth muscle to the point where it may be a slight challenge to distinguish between smooth muscle, fibroblasts, and carcinoma cells. And then the cells may even be denser at the serosal surface.” The space between tumor cells can make the specimen appear much less cellular than a solid tumor involving the same area. But tumor cells may blend into the background, so the specimen may be more cellular than it appears at first glance.

The combined positive score is appropriate for analyzing metastatic gastric cancer to a lymph node, Dr. Hanks says, but the magnification level is important. “Obviously, lymph nodes can have a low level of PD-L1 positivity in adjacent normal lymphoid tissue, so the 20x field recommendation is applied in lymph node.” Agilent’s E-Learning module provides examples of this application of the scoring method.

Scoring strategies provided by Agilent also include approaches that can be used with one patch or multiple patches of positivity, Dr. Hanks notes. “If you have a sea of negativity and one small patch that’s a CPS of one, obviously your overall score is CPS zero or negative—no PD-L1 expression.”

Fig. 6. CPS: Example calculation method tumor area with staining



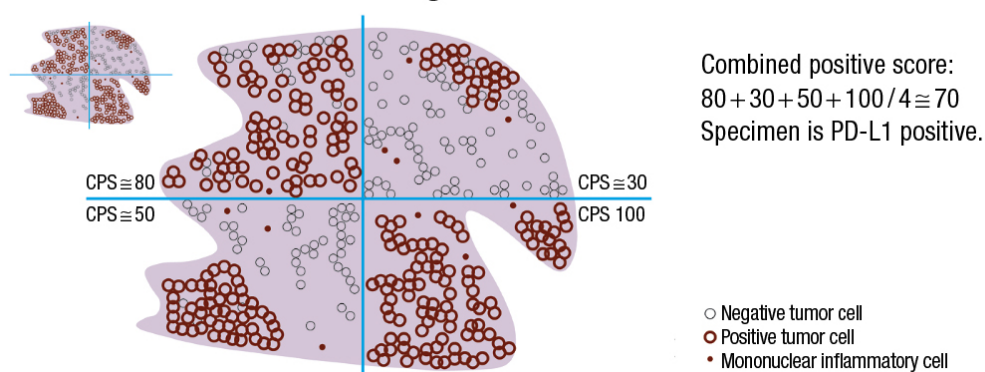
“If you have, as in this case (**Fig. 6**), a point where there is a patch where the CPS is actually 80, and all of the rest of it is negative and it represents 10 percent of your specimen, the math works out that the CPS is eight, so this specimen is PD-L1 positive.”

Agilent offers a “divide and conquer” strategy for scoring large resection specimens, which involves dividing a large section of tissue into four equal quadrants and solving by sector, Dr. Hanks says. “You would divide the tissue up so there are approximately an equal number of cells in each quadrant, then add up the scores and divide by four.” (**Fig. 7**).

A common mistake is to underestimate the denominator in the CPS by underestimating the number of unstained tumor cells, resulting in a falsely elevated CPS and possibly a mistaken eligibility for anti-PD-1 treatment. Similarly, if the denominator is overcounted, the patient is at risk of not qualifying for treatment. The only way to avoid such errors is to practice the method, Dr. Hanks emphasizes.

Combined positive scores are reliable, especially around the critical cutoff of one, in determining PD-L1 expression or no PD-L1 expression, she says. Agilent has found, in analyzing scores across instruments, platforms, and pathologists, that consistency is also high. “CPS in gastric cancer passed all statistical criteria for solid reproducibility and intra-observer variability. After training and practice, pathologists can align with this algorithm

Fig. 7. CPS: Example calculation method heterogeneous tumor area



and can reproducibly and reliably score gastric cancer specimens at the CPS ≥ 1 ."

As a general precept, "we recommend that evaluation of PD-L1 stains be performed within the context of the pathologist's past experience and best judgment in interpreting IHC stains." However, the most difficult part of a diagnostic workup using the PD-L1 assay in gastric cancer is the CPS, she believes. "It takes time to see a few cases to get a low-power gestalt. Are you scoring something that's obviously positive with a high CPS versus obviously a zero? Those are the easy cases."

Since each case is different, each may require a different strategy, and different pathologists may prefer one method over another, Dr. Hanks points out. "We're finding that pathologists have different ways to approach cases and they find what works for them." But some of her advice applies to all pathologists who wish to avoid problems with complex cases. "Make sure you see all of the positive and all of the negative. Have you considered the denominator? After doing your combined positive score calculation, review the slide. Does it make sense? Can you defend your score? Can you reproduce your score? And practice, practice, practice."

As with any companion diagnostic with a cutoff score, more time is spent around the cutoff. "In my experience, when I am near that CPS of one with a gastric cancer case, whether it's intestinal or diffuse pattern, I know I'm going to have to spend a little more time. It may take me 10 minutes to sit and make sure I've looked at all the positives and all the negatives to have an estimation of what the denominator is before I make the decision."

How results are reported will vary depending on the laboratory's information system, Dr. Hanks notes. An example of how to report results is included in the interpretation manual and in principle four of the first E-Learning module. "The FDA required that our example include how many biopsies were actually obtained and have that in the report. In addition," she says, "the report shows the CPS score and whether the sample result is interpreted with PD-L1 expression or no PD-L1 expression."

"There is a full page of other details that we suggest should be included." Among these are the type of tissue, number of biopsies in the tissue block, and check-offs covering control cell line slide results and the adequacy of tumor cells present. However, Dr. Hanks says, "Including a lot of this information is really the final decision of the laboratory director or the medical director of the lab."

In response to questions about use of different platforms to bring Agilent's PD-L1 IHC test online, Dr. Eklund cautions that Agilent does not endorse off-label use of its products. "If users are modifying any part of the validated full solution approved by the FDA, it represents a laboratory-developed test, for which users are responsible for conducting a full validation according to country-specific recommendations. Different organizations globally have published recommendations on validation of IVD devices, but ultimately it is the responsibility of the medical director in the laboratory to ensure that the tests they are using are fit for the purpose."

The CPS process is not only valuable but also relatively easy to learn, Dr. Hanks says. Agilent conducted training in a pre-commercial setting and found that pathologists in private practice who were trained performed just as well at scoring the CPS as they did with the tumor proportion score.

"Pathologists may wonder if this can be applied in their practice and I think it's important for them to take time to look at the E-Learning modules and get practice with the algorithms. Pathologists whom we have trained catch on to scoring with CPS very quickly, and even those who haven't scored using TPS in lung cancer in the past also catch on. It's just a matter of trying it—and practicing."

[hr]

Anne Paxton is a writer and attorney in Seattle. The Agilent E-Learning modules are at www.agilent.com/en-us/e-learning-dako-products. Log-in is required.