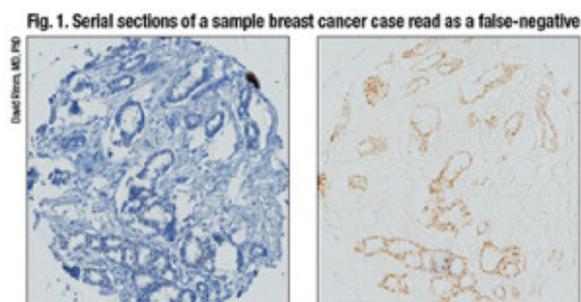


PD-L1, other targeted therapies await more standardized IHC

Anne Paxton

February 2016—Immunohistochemistry is heading down a path toward more standardization, and that's essential as it plays an increasing role in rapidly expanding immunotherapy, says David L. Rimm, MD, PhD, professor of pathology and of medicine (oncology) and director of translational pathology at Yale University School of Medicine. As a co-presenter of a webinar produced by CAP TODAY in collaboration with Horizon Diagnostics, titled "Immunohistochemistry Through the Lens of Companion Diagnostics" (http://j.mp/ihc_lens_webinar), he analyzes the core challenges of IHC's adaptation to the needs of precision medicine: binary versus continuous IHC, measuring as opposed to counting or viewing by the pathologist, automation, and assay performance versus protein measurement.

"Immunohistochemistry is 99 percent binary already," Dr. Rimm points out. "There are only a few assays in our labs—ER, PR, HER2, Ki-67, and maybe a few more—where we really are looking at a continuous curve or a level of expression."



The left panel shows a case that was called negative in one lab and positive in another. The right panel is a serial section of that case showing definitive positive staining illustrated by omission of hematoxylin counterstain.

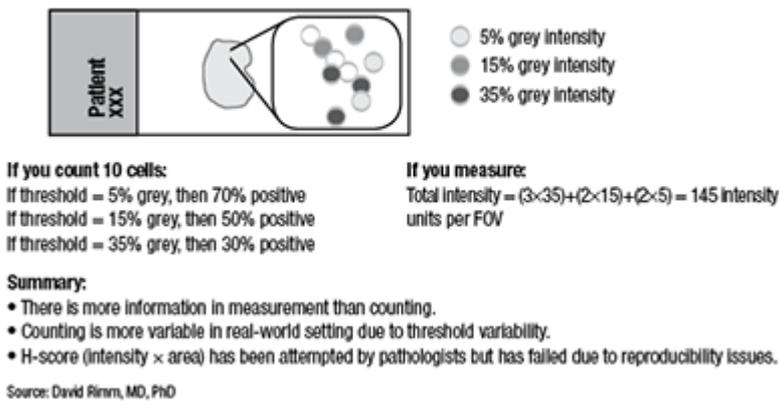
Two criteria in the 2010 ASCO/CAP guidelines on ER and PR testing in breast cancer patients are key, he says: 1) the percentage of cells staining and 2) any immunoreactivity. "The first is hard to estimate, but the guidelines recommend the use of greater than or equal to one percent of cells that are immunoreactive. That means they could have a tiny bit of signal or they could have a huge amount of signal and they would be considered immunoreactive, which thereby makes this a binary test."

Having the test be binary can be a problem for companion diagnostic purposes because any immunoreactivity is dependent on the laboratory threshold and counterstain. For example, if two of the same spots, serial sections on a tissue microarray, were shown side by side, one with and one without the hematoxylin counterstain, "you might see the counterstain make this positive test into a negative by eye, which is a potential problem with IHC when you have a binary stain." (**Fig. 1**).

Dr. Rimm describes a small study done with three different CLIA-certified labs, each using a different FDA-approved antibody and measuring about 500 breast cancer cases on a tissue microarray. The study showed there can be fairly significant discordance between labs—between 18 and 30 percent discordance—in terms of the cases that were positive. "In fact, if we look at outcome, 18 percent of the cases were called positive in Lab Two but were negative in Lab Three. Lab Three showed outcomes similar to the double positives whereas Lab Two had false-

negatives.” This is an important problem that occurs when we try to binarize our immunohistochemistry, he says.

Fig. 2. Counting vs. measuring



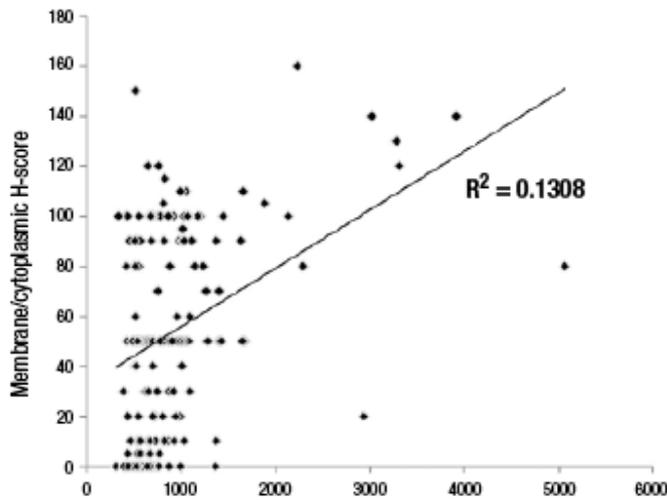
Counting is more variable in a real-world setting due to the variability of the threshold for considering a case positive. “You can easily calculate that if your threshold was five percent, then you’d have 70 percent positive cells. And you would easily call this positive. But if you added more hematoxylin because that’s how your pathologist liked it, then perhaps you’d only have 30 percent positive. So this is the risk of using thresholds.” (Fig. 2).

Although this is done in all of immunohistochemistry today, Dr. Rimm thinks it is an important consideration as IHC transitions to more standardized form. “An H score—intensity times area, which has been attempted many times, can’t be done by human beings. Pathologists try but have failed.”

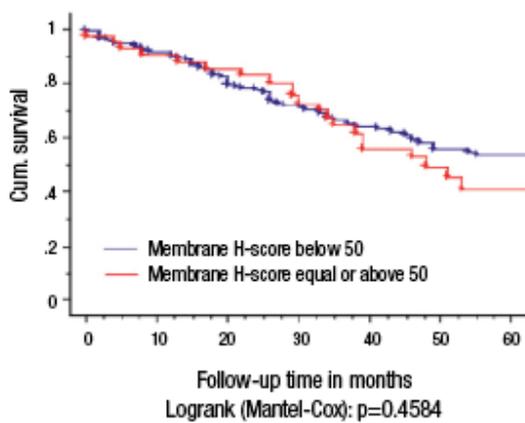
“We can’t do those intensities by eye. We have to measure them with a machine. But we get a very different piece of information content when we measure intensity, as opposed to measuring the percentage of cells above a threshold. In sum, more information is present in a measurement than in counting.”

Fig. 3. Example of H-score failure

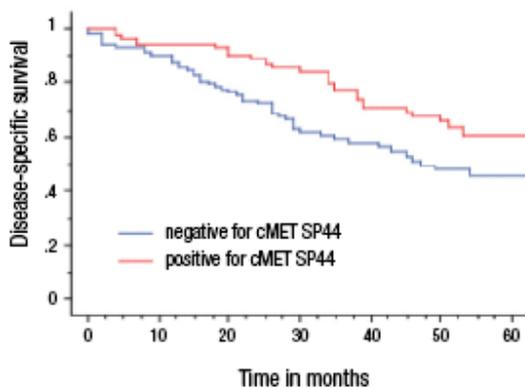
A MET AQUA vs. DAB IHC scoring by eye



B Kaplan-Meier cumulative survival plot for total membrane protein-score binarized at 50



C Kaplan-Meier for YTMA 250-2; NSCLC stratified for cMet expression, 5-year follow-up



A shows comparison of a quantitative fluorescence score on the x axis versus an H-score on the y axis. Note the noncontinuous nature of human estimation of intensity times area (H-score). B) The survival curve in a population of lung cancer cases using the H-score. C) The survival curve in the same population using the

quantitative score. (Source: David Rimm, MD, PhD)

Pathologists read slides for a living, so it's uncomfortable to think about giving that up in order to use a machine to measure the slides. "But I think if we want to serve our clients and our patients, we really owe them the accuracy of the 21st century as opposed to the methods of the 20th century." (**Fig. 3**).

Among the currently available quantitative measuring devices are the Visiopharm, VIAS (Ventana), Aperio (Leica), InForm (Perkin-Elmer), and Definiens platforms. "We use the platform invented in my lab, called Aqua [Automated Quantitative Analysis], but this is now owned by Genoptix/Novartis. Genoptix intends to provide commercial tests using Aqua internally," Dr. Rimm says, "as well as enable platform and commercial testing through partnership with additional reference lab providers.

"There are many quantification platforms," he adds, "and I believe that any of them, used properly, can be effective in measurement."

(Of the 265 participants in the CAP PM2 Survey, 2015 B mailing, who reported using an imaging system for quantification, 4.6 percent use VIAS, 4.1 percent use ACIS, 0.8 use Applied Imaging, and 10 percent use "other" imaging systems. Of the 1,359 Survey participants who responded to the question about use of an imaging system to analyze hormone receptor slides, 1,094, or 80.5 percent, reported not using any imaging system for quantification.)

Says Dr. Rimm: "The first platform we used to try to quantitate some DAB stain slides was actually the Aperio Nuclear Image Analysis algorithm. But the problem with DAB is that you can't see through it. And so inherently it's physically flawed as a method for accurate measurement." He compares DAB to looking at stacks of pennies from above, where their height and quantity can't be surmised, as opposed to from the side, where their numbers can be accurately estimated. "This is why I don't use, in general, DAB-type technologies or any chromogen."

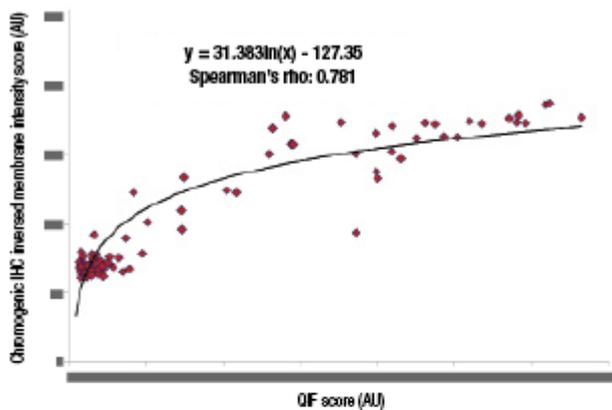
Fluorescence doesn't have this problem, and that is the reason Dr. Rimm began using fluorescence as a quantitative method. "We try to be entirely quantitative without any feature extraction. So we define epithelial tumors using a mask of cytokeratin. We define a mask by bleeding and dilating, filling some holes, and then ultimately measure the intensity of each cell, or of each target we're looking for. In this case, in a molecularly defined compartment."

Compartments can be defined by any type of molecular interactions. "We defined DAPI-positive pixels as nuclei, and we measure the intensity of the estrogen receptor within the compartment. And that gives us an intensity over an area or the equivalent of a concentration." Many other fluorescent tools can be used in this same manner, but he cautions against use of fluorescent tools that group and count. "That's a second approach that can be used, but the result gives you a count instead of a measurement."

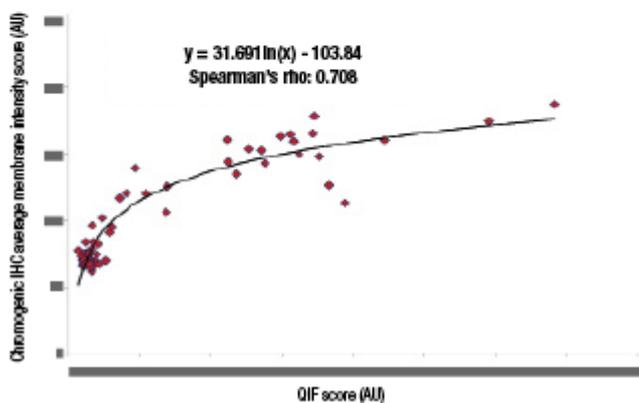
When comparing a pathologist's reading versus a quantitative immunofluorescence score, he notes, pathologists actually don't generate a continuous score. Instead, pathologists tend to use groups. "We tend to use a 100 or a 200 or an even number. We never say, 'Well, it's 37 percent positive.' We say, 'It's 40 percent positive,' because we know we can't reproducibly tell 37 from 38 from 40 percent positive."

Fig. 4. Comparison of QIF score with chromogenic score shows decreased dynamic range of the chromogenic method due to assay saturation

CB11 Regression of HER2 CB11 chromogenic IHC and QIF scores in YTMA 263



SP3 Regression of HER2 SP3 IHC vs. QIF scores in YTMA 263



Source: Daniel Carvajal-Hausdorf, MD

The result of that is a noncontinuous scoring result, which doesn't give the information content of quantitative measurement. A comparison between the two methods shows that at times, where quantitative measurement shows a significant difference in outcome, nonquantitative measure or an H-score difference may not show a difference in outcome. (Fig. 3 illustrates this concept.)

"Pathologists tend to group things, and we also tend to overestimate. It's not that pathologists are bad readers. It's just the tendency of the human eye because of our ability to distinguish different intensities and the subtle difference between intensities. But even if you compare two quantitative methods, you can see that the method where light absorbance occurs—that is the percent positive nuclei by Aperio, which is a chromogen-based method—tends to saturate. This is, in fact, amplified dramatically when you look at something with a wide dynamic range like HER2." (Fig. 4).

In one study, researchers found less than one percent discordance—essentially no discordance—between two antibodies (Dekker TJ, et al. *Breast Cancer Res.* 2012;14[3]:R93). But looking at these results graphed quantitatively, you would see a very different result, Dr. Rimm says. "You can see a whole group of cases down below where there's very low extracellular domain and very high cytoplasmic domain. In fact, some of these cases have essentially no extracellular domain, but high levels of cytoplasmic domain, and other cases have roughly equal levels of each" (Carvajal-Hausdorf DE, et al. *J Natl Cancer Inst.* 2015;107[8]:pii:djv136).

Recent studies by Dr. Rimm's group have shown this to have clinical implications. He looked at patients treated with trastuzumab in the absence of chemotherapy, in an unusual study called the HeCOG (Hellenic Cooperative Oncology Group) trial.

"We found that patients who had high levels of both extracellular and intracellular domain have much more benefit than patients who are missing the extracellular domain and thereby missing the trastuzumab binding site." Follow-up studies are being done to validate this finding in larger cohorts.

Preanalytical variables, Dr. Rimm emphasizes, can have significant effects on IHC results, and more than 175 of them have been identified. "These are basically all the things we can't control, which is the ultimate argument for standardization."

In a surprising study by Flory Nkoy, et al., he says, it was shown that breast cancer specimens were more likely to be ER negative if the patient’s surgery was on a Friday because there was a higher ER-negative rate on Friday than on Monday. “So how could that be? Well, it was clearly the fact that the tissue was sitting over the weekend. And when it sat over the weekend, the ER positivity rate was going down” (*Arch Pathol Lab Med.* 2010;134:606-612).

Another study showed that after one hour, four hours, and eight hours of storage at room temperature, you lose significant amounts of staining, Dr. Rimm says. “And perhaps the best nonquantitative study or H-score-based study of this phenomenon was done by Isil Yildiz-Aktas, et al., where a significant decrease in the estrogen receptor score was found after only three hours in delay to fixation” (*Mod Pathol.* 2012;25:1098-1105).

How long the slide is left to sit after it is cut is another preanalytical variable to be concerned with. “In the clinical lab, that’s not often a problem since we cut them, then stain them right away. But in a research setting, a fresh-cut slide can look very different from a slide that’s two days old, six days old, or 30 days old, where a 2+ spot on a breast cancer patient becomes negative after 30 days sitting on a lab bench. So those are both key variables to be mindful of.”

One solution for those preanalytic variables is trying to prevent delayed time to fixation. “And probably time to fixation is one of the main preanalytic variables, although it’s only one of the many hundreds of variables. The method we use to try to get around this problem is to use core biopsies or allow rapid and complete fixation, and then other things can be done.”

Finally, he warns, don’t cut your tissue until right before you stain it. “If you’re asked to send a tissue out to a collaborator or someone who is going to use it for research purposes later, we recommend coring and re-embedding the core, or sending the whole block. Unstained sections, when not properly stored in a vacuum, will ultimately be damaged by hydration or oxidation, both of which lead to loss of antigenicity.”

The crux of the matter is assay performance versus protein measurement, Dr. Rimm says. “In the last six to nine months, we really are faced with this problem in spades, as PD-L1 has become a very important companion diagnostic.”

There are now four PD-L1 drugs with complementary or companion diagnostic tests (**Fig. 5**). One of the FDA-approved drugs, nivolumab (Opdivo, Bristol-Myers Squibb), for example, uses a clone called 28-8, which is provided by Dako in an assay, a complementary diagnostic assay, and with the following suggested scoring system: one percent, five percent, or 10 percent. In contrast, pembrolizumab (Keytruda, Merck) is also now FDA-approved but requires a companion diagnostic test that uses a different antibody, although the same Dako Link 48 platform. This diagnostic has a different scoring system of less than one percent, one to 49 percent, and 50 percent and over.

Fig. 5. Companion Dx assays for PD-L1

PD-L1 Drug and Vendor	Nivolumab BMS	Pembrolizumab Merck	Atezolizumab Roche/Genentech	Durvalumab AstraZeneca
Clone and source and epitope (ECD=extra-cellular domain; ICD=intra-cellular domain)	28-8 Abcam – ECD	22c3 Dako – ECD	SP142 Spring Bio – ICD	SP263 Spring Bio – ECD
IVD class III partner	Dako	Dako	Ventana	Ventana
Scoring method	% cells with membrane staining at any intensity	% cells with membrane staining at any intensity	TC=Tumor cell IC=Immune cell Combine both percentage and subjective intensity	% cells with membrane staining at any intensity
Thresholds	>1% >5% >10%	<1% 1–49% >50%	TC3=TC>50% IC3=IC>10% TC2/IC2=TC or IC>5% TC1/IC1=TC or IC>1%	>25%
Method	Pathologist/ Subjective	Pathologist/ Subjective	Pathologist/ Subjective	Pathologist/ Subjective

Source: David Rimm, MD, PhD

Two other companies, Roche/Genentech and AstraZeneca, also have drugs in trials that may or may not have

companion diagnostic testing, though both have already identified a partner and a unique antibody (neither of those listed above) and companion diagnostic testing scores used in their clinical trials.

“So what’s a pathologist to do?” Dr. Rimm says. “Well, there are a few problems with this. First of all, what we really should be doing is measuring PD-L1. That’s the target and that’s what should ultimately predict response. But instead what we’re stuck with, through the intricacies of the way our field has grown and our legacy, is closed-system assays. While these probably do measure PD-L1, we do not know how these compare to each other.” Two parallel large multi-institutional studies are addressing this issue now, he says.

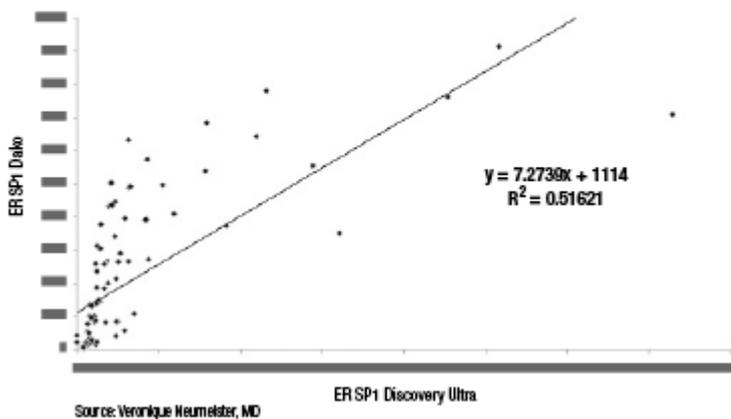
There are solutions for managing these closed-system assays to be sure the assay is working in your lab and that you can get the right answer, Dr. Rimm says. His laboratory uses a closed-system assay for PD-L1, relying not on the defined system but rather on a test system it has developed in doing a study with different investigators.

Sample runs by these different investigators show the potentially high variability, he says. “In a scan of results, no one would deny which spots are the positive spots and which are the negative.” But the difference in staining prevents accurate measurement of these things and shows the variability inherent even in a closed-box system.

A comparison of two closed-box systems, the SP1 run on the Discovery Ultra on Ventana, and the SP1, same antibody, run on the Dako closed-box system, also shows that, in fact, there’s not 100 percent agreement using same-day, same-FDA-cleared antibody staining and different autostainers. So automation may not solve the problem, Dr. Rimm notes (**Fig. 6**).

Fig. 6. Automation may not solve the problem: same day, same FDA-cleared antibody staining; different autostainers quantified using Aperio pixel counter software

ER SP1 on YTMMA 201; same day staining on Discovery Ultra at 1:30 dilution and Dako autostainer at 1:100 for 1 hour



“When running these in a quantitative fashion and measuring them quantitatively, there are actually differences in the way these closed-box systems run. And so you, as the pathologist, have to be the one who makes sure your assays are correct, your thresholds are correct, and your measurements are accurate.”

The way to do that, he believes, is to use standardization or index arrays. An index array of HER2 that his laboratory developed has 3+ amplified, 2+ amplified, not amplified, and so on from 80 cases in the lab’s archive, shown stained with immunofluorescence and quantitative and DAB stain. “It was only with this standardization array, run every time we ran our stainer, that we were able to draw the conclusions in the previous study about extracellular versus cytoplasmic domain.”

Companies have realized the importance of this, and specifically companies like NantOmics (formerly OncoPlexDx) have realized they can exactly quantitate the amount of tissue on a slide using a specialized mass spectrometry method, he says. “They can actually give you amol/μg of total protein.”

He and colleagues are working with NantOmics now to try to convert from amols to protein to average quantitative fluorescent scores to help build these standards and make standard arrays more accurate. “This is still a work in progress, but I believe this is ultimately the kind of accuracy that can standardize all of our labs. We have shown that the quantitative fluorescence system is truly linear and quantitative for EGFR measurements when using mass spectrometry as a gold standard.” They are preparing to submit a manuscript with this data.

In the interim, Dr. Rimm's laboratory has begun working also with Horizon Diagnostics, employing Horizon's experimental 15-spot positive-control array. "When you use this array and quantitate it with quantitative fluorescence, you get a very interesting profile. If a cut point is set at one point, you would see three clearly positive cells or spots and 12 clearly negative spots with two different antibodies. But is that the threshold?"

"In fact, using a little higher score and a very quantitative test, you might find that the threshold may, in fact, be a little bit lower than that." It turns out that only three of these 12 spots are true negatives. The others at least have some level of RNA, and some have a lot. "So how do we handle these? And are these behaving the same way with multiple antibodies?" Parallel results, finding nearly the same threshold case, have been found using SP142 from Ventana, E1L3N from Cell Signaling, and SP263 from Ventana.

Studies to address those issues are still in the early stage, he says. He cautions that there is variance in these assays, and more work is being done to reproduce the data. "But I think the important point is that, using these kinds of arrays, you can definitively determine whether your lab has the same cut point as every other lab. And were we to quantitate this with mass spectrometry, we would know exactly the break point for use in the future."

Dr. Rimm's laboratory has also built its own PD-L1 index tissue microarray with a number of its own tumor slides ranging from very low to very high expressors, a series of cell lines, and including some placenta-positive controls on normal tumor. He has found that generating an index array has advantages, and he encourages other laboratories to prepare their own index arrays to increase the accuracy and reproducibility of their laboratory-developed tests. "You can produce these in your own lab so that you can be sure you can standardize your tests run in your clinical lab from day to day and week to week as part of an LDT."

"If we think about it, there really are no clinical antibodies today that are truly quantitative," Dr. Rimm says. "And when there are, new protocols will be required, but I believe those protocols are now in existence. We just await the clinical trials that require truly quantitative protein measurement or in situ proteomics."

In that process of moving toward in situ proteomics, suggests web-inar co-presenter Clive Taylor, MD, DPhil, professor of pathology in the Keck School of Medicine at the University of Southern California, FDA approval, per se, will not solve any of the problems discussed in the webinar. (See the January 2016 issue for the full report of Dr. Taylor's presentation.) "I think what the FDA approval will do is demand that we find solutions to these problems ourselves. The FDA's attitude is, to a large degree, dependent on the claim. So if we just use immunohistochemistry as a simple stain, then the FDA classes that as sort of class I, level 1. And we can do that [IHC stain] without having to get preapproval by the FDA.

"On the other hand, if we take something like the well-established HercepTest, where based on the result of that test alone, it's decided whether or not the patient gets treatment, treatment that's very expensive and treatment that has benefits and...side effects. That claim is, in fact, a very high-level claim. And for that, the FDA is demanding high-level data, which I think is entirely appropriate," Dr. Taylor says.

Most of these upcoming companion diagnostics, if not all, he says, will be regarded by the FDA as class III, high level or high complexity. They will require a premarket approval study in conjunction with a clinical trial. And the FDA will demand high standards of control and performance, eventually. "There are not many labs that can produce those high standards as in-house or lab-developed tests today. And even the companies currently in trials are not producing the improved performance level for these tests that we are talking about today, as being required for high-quality quantitative and reproducible companion diagnostics. Eventually, I am convinced we will have to do that. It's just that it will take time to get there."

The FDA can only approve what is brought to it, Dr. Rimm points out. And so a true, fully quantitative IHC-based assay has presumably never been submitted, or at least never been approved by the FDA. "What we're seeing instead are the assays that the FDA has approved, which are well defined and rigorously submitted. However, the result is a closed system that we use, which may or may not accurately measure PD-L1 on the slide, depending upon preanalytic variables and individual laboratories' methods."

“So questions keep popping up. And I can only say that we, as pathologists, have the final responsibility to our patients. And while it may not be recommended and it may change in the future, right now lab-derived tests or LDTs may be more accurate than FDA-approved platforms.”

“If you think about it, in molecular diagnostics where I’m familiar with EGFR and BRAF and KRAS tests, in that testing setting, less than 25 percent of the labs that do that test actually use the FDA-approved test,” Dr. Rimm says. “The remainder of the labs do their own LDTs, including our labs here at Yale.”

It wouldn’t surprise him if the same thing happens for PD-L1. “I’m aware of at least two labs—and we probably will be the third—that devise our own LDT for PD-L1 testing using the standards I’ve discussed, using array-type controls to be sure that our levels are correct, and then using a scoring system that we derived.”

“We aren’t really in a position to know at the time that we receive a piece of lung cancer tissue whether the oncologist is going to use pembrolizumab, which requires a companion diagnostic, or nivolumab, or the other drugs, which may or may not require a companion diagnostic. So in that sense, we’re almost bound to use an LDT,” Dr. Rimm says, since his lab can’t actually run four different potentially incongruent, though FDA-approved, tests for PD-L1.

Until a truly quantitative approach is developed and submitted to the FDA and approved, Dr. Taylor believes we won’t see things changing. “The algorithms that currently are approved have been approved on the basis that they can produce a similar result to a consensus group of pathologists. So they’re only as good as the pathologists.”

“As Dr. Rimm has discussed, I actually believe we can get a much better result than the pathologists can get with their naked eye. We have to get away from comparing it to what we currently can do and start to try to construct a proper test, just like we did in the clinical lab 30 years ago when we automated the clinical lab,” Dr. Taylor says. “We need to automate anatomic pathology, including the sample preparation, the assay process, and the reading, all three together in a closed system. And we’re nibbling away at the edges of it. We’ll get there, but it’ll take some time.”

Dr. Rimm is skeptical that the diagnostics field has learned any lessons from HercepTest and the companion diagnostics world of almost 20 years ago. “The submissions to the FDA for PD-L1 look very similar to what was submitted in 1998 for the HercepTest, the companion diagnostic test for trastuzumab [Herceptin]. And that’s disappointing. I think that is 20-year-old technology and we can do better. But even if we want to use the 20- or 40-year-old DAB-based technology, we should still be standardizing it and having a mechanism for standardization and having defined thresholds.”

As future FDA submissions come in, Dr. Rimm hopes that “even if they’re not quantitated, they can be standardized as to where the thresholds occur, so that we can be sure we deliver the best possible care to patients. And in the interim, I think we, as pathologists, will have to do that standardization with an LDT to be sure we’re giving our best results.”

Dr. Taylor warns that there is only a limited number of labs in the country and in the world that will be able to produce these LDTs, because of the complexity. “The FDA has already said in a position paper that it believes it may have to regulate LDTs to some extent. And what that will mean is that in the validation process, your own LDT will start to approach what is required for an FDA-approved test. And most labs are in no position to be able to do that.”

“So I think we’re going to come to a blending here, all forced by companion diagnostics. This is in situ proteomics,” Dr. Taylor says. “It’s a new test, essentially. It’s not straightforward immunohistochemistry, but a new test. And I think the fluorescence approach that Dr. Rimm has used has a lot of advantages in relating signal to target in terms of figure out what the best test is and stop comparing it to the pathologists. We should compare it to the best assay we can produce.”

With respect to the PD-L1 problem, Dr. Rimm notes, “I would point out that there is a so-called ‘Blueprint’ for

comparison of the different antibodies and the different FDA assays, or potentially FDA-submitted tests anyway, to see how equivalent they are.” Similarly, he adds, the National Comprehensive Cancer Network recently issued a press release describing a multi-institutional study to assess the FDA-approved assay but also including an LDT (the Cell Signaling antibody E1L3N using the Leica Bond staining platform).

He points to a newly published study by his group (McLaughlin J, et al. *JAMA Oncol.* 2016;2[1]:46-54), finding that objective determination of PD-L1 protein levels in non-small cell lung cancer reveals heterogeneity within tumors and prominent interassay variability or discordance. The authors concluded that future studies measuring PD-L1 quantitatively in patients treated with anti-PD-1 and anti PD-L1 therapies may better address the prognostic or predictive value of these biomarkers. With future rigorous studies, including tissues with known responses to anti-PD-1 and anti-PD-L1 therapies, researchers could determine the optimal assay, PD-L1 antibody, and the best cut point for PD-L1 positivity.

Other work that will probably come out in mid-2016 from Dr. Rimm’s group has shown that expression of PD-L1 is largely bimodal, he says. “That is, there’s a group of patients that express a lot, and then there’s another group of patients that expresses a little or none.”

So time will tell how PD-L1 will be scored. “But if you look at the data from the Merck study and their cut point of greater than 50 percent, or even the cut point from the AstraZeneca studies of greater than 25 percent, you’re really dichotomizing the population into patients who are truly PD-L1 positive from patients who are negative or almost negative.”

“Of course, we don’t want to miss patients in that negative to almost-negative group who will respond,” Dr. Rimm says. “On the other hand, we probably will have fairly good specificity and sensitivity with the assay defined by Merck and Dako with 22C3 as was recently published” (Robert C, et al. *N Engl J Med.* 2015;372[26]:2521-2532).

Many difficulties lie ahead, as researchers try to weigh the merits of different drugs with different approved tests on different platforms, involving different antibodies, Dr. Taylor says. “Does the lab try to set up four different PD-L1s, and if we only have one platform and not another, what do we do about that?” He thinks the tests may often be sent out to larger reference labs or academic centers as a result.

Dr. Rimm confirms that his own lab’s LDT—although literally thousands of PD-L1 tests have been conducted using it—is not yet up and running in the Yale CLIA laboratory, and in the meantime the IHC slides are being sent out to a commercial vendor.

Eventually, Dr. Taylor believes, the pressure of these dilemmas will lead the diagnostics field to develop an immunoassay on tissue sections. “We’ve never been forced to do that before, but once we are, that will produce a huge change in diagnostic capability and research capability.”□

Anne Paxton is a writer in Seattle.