

Gastric HER2, hsALK to join monitored PT list

Anne Paxton

September 2021—Beginning next year, CAP-accredited laboratories that perform HER2 immunohistochemistry in gastroesophageal adenocarcinoma or highly sensitive (hs) ALK in non-small cell lung cancer will be required to enroll in proficiency testing for those analytes. This change comes as the first new IHC predictive biomarker assays are to be added to the monitored list in a decade.

The new requirements are notable not only because the last year an IHC biomarker became monitored was 2011, when ER/PgR for breast cancer joined the list, but also because the requirements culminate many years of consideration and conversation by CAP Immunohistochemistry Committee and CAP Accreditation Program leaders. Reflecting the complexity of the decisions to make these two biomarker tests monitored, the committee has released a list of frequently asked questions to help laboratories comply with the proficiency testing requirement and to improve their testing processes (<https://bit.ly/CAP-IHC-FAQ>).

Andrew Bellizzi, MD, chair of the CAP Immunohistochemistry Committee, was an advocate for an expansion of the proficiency testing requirements to include non-breast predictive markers. “Probably the most important pillar of quality is the peer laboratory inspection process. But hand in hand with that is proficiency testing,” says Dr. Bellizzi, director of immunohistochemistry and GI pathology at the University of Iowa Hospitals and Clinics.



Dr. Bellizzi

“The main problems in quality in predictive marker IHC reside in suboptimal IHC protocols,” Dr. Bellizzi says, explaining that the decision to proceed with required proficiency testing was motivated in part by the need to address that quality problem. “PT for the highly sensitive ALK assay was where we had the best data. And we consistently had around 15 percent of labs failing that IHC Survey. But it’s not required, so there is no enforcement mechanism.”

Predictive markers are often a tiny fraction of the analytes in a given IHC laboratory but their results carry the greatest weight, with a positive or negative indicating whether a patient may benefit from a specific therapy, Dr. Bellizzi notes. “If the biomarker is negative, that result cannot be predicted based on morphology, so there’s no backup. There’s no crutch. That is unlike typical diagnostic IHC markers, which are used in panels and in combination with morphology and a lot of clinical information to arrive at an overall interpretation of the case.”

The IHC laboratory resides in anatomic pathology, which historically has been an inherently qualitative discipline compared with clinical pathology—the latter with its emphasis on objective metrics, precision, and reference standards, says Dr. Bellizzi. Proficiency testing requirements set by the Centers for Medicare and Medicaid Services appear to reflect this distinction. “There are 81 analytes for which participation in proficiency testing is required by CMS. They are CMS-regulated analytes. They’re all very important tests like Gram stain and CBC and calcium and albumin, but they are all clinical pathology tests.”

“When I joined the IHC Committee nearly a decade ago,” he says, “and I started leafing through the CAP PT catalog, honestly I was shocked that breast HER2 and ER were not CMS-regulated for proficiency testing. Of all the tests in pathology, results of predictive marker IHC have among the greatest clinical consequence.” A decade ago, he adds, the CAP made up for this gap by monitoring the breast biomarkers.

There is the view, Dr. Bellizzi says, that the art and science of medicine could be implicated in proficiency testing monitoring decisions, but he sees it differently. He points out that the IHC readout is part of the analytical phase of the test and whether a HER2 slide is zero or 1+, 2+, or 3+ should be an objective truth. “The result is not being spit out automatically like a chemistry test,” he says, but “when we’re reading predictive marker IHC, to some extent we’re functioning like a chemistry analyzer. There’s no room for personal opinion. What we’re targeting by proficiency testing for these predictive markers is the actual IHC protocol itself.”

Although Dr. Bellizzi describes himself as an “IHC guy,” he considers molecular and IHC techniques to be entirely complementary. “It depends on the specific diagnostic or predictive context. In some instances the molecular is the best. It’s inherently better at multiplexing so you’re able to assess more analytes simultaneously. Most IHC assays are singleplex or dualplex, while for one molecular test you might get 500 answers. But also, with the extraction, the testing, the computational stuff, and the analysis, the turnaround time may be a week or two weeks.”

IHC is fast, says Emily Meserve, MD, MPH, IHC Committee member and technical consultant for immunohistochemistry at NorDx Laboratories, Scarborough, Me., and staff pathologist for Spectrum Healthcare Partners at Maine Medical Center. “In most labs, you can get results within 24 hours,” and they can often be obtained in a few hours or half a day. “And in most cases the test can be interpreted pretty easily by a subspecialty pathologist or a general surgical pathologist.” The typical price for an IHC assay is in the hundreds of dollars while molecular can cost \$1,000 or more, she adds, depending on how complex the assay is.

BRAF is a good example of a biomarker for which the decision to use IHC or a molecular assay is context dependent, Dr. Meserve says. “We have a mutation-specific IHC stain that only identifies a very specific mutation in the *BRAF* gene, but there are other mutations in the gene that are important to identify. So if the immunostain is positive, you’ve confirmed mutation, but if it’s negative, you haven’t fully examined *BRAF* so you may still need to do molecular. At this point, some labs say, ‘I don’t want to waste my time with the immunostain; I’m going to go straight to molecular.’”

When Dr. Meserve became an IHC Committee member last year, the discussion over requiring PT testing for the HER2 gastric and highly sensitive ALK lung biomarkers had already been underway for several years. She has found PT is important because it provides an opportunity to do an assessment outside one’s own laboratory. “I spend a lot of time counting ER-positive breast cancer cases and determining percent positives in my patient population, and I can compare that to published literature to see if I’m detecting about the right number. But a far superior method is to compare directly to the results of the same tumor tested at hundreds of laboratories across the country.”

In addition, requiring PT for highly sensitive ALK in NSCLC and for gastric HER2 “might help us realize there are more laboratories in a gray zone, that are actually doing pretty well except for certain situations, for which their assays may only need minor adjustment. And that’s better for everybody.”

Often, “you don’t know what you don’t know,” Dr. Meserve says. She cites one case study in which a laboratory performing ER by IHC encountered problems because its assay was not calibrated correctly. “PT is how labs find this out. They thought they were doing everything just fine until they started comparing their results to other labs.” But “the vast majority of labs in this country are doing the right things and the assays are performing very well. Major issues with assay performance are relatively infrequent.”

“What Dr. Bellizzi did in aggregating the data,” she says, “was point out that, yes, it’s a small number but it’s not zero. And patients will benefit if we address this.” CAP leaders agreed.

Just as therapy-related decisions are made for patients with breast cancer based on the results of ER and HER2, she notes, the same weight is put on gastric HER2 in patients with gastroesophageal adenocarcinoma and lung cancer when they have a mutation. “And we should make sure we’re monitoring these carefully, too, and holding laboratories to a high standard.”

This can sometimes be an issue, she says, based on the feeling that anatomic pathology is qualitative, not quantitative, and its values should not be assessed the same way as a quantitative test. “I agree there is an art and science involved in interpretation, but I fall on the side of thinking the standard is helpful and useful.” Her training in public health and epidemiology has tended to convince her that “at some point it is beneficial to regulate at a large level so that the care for everybody is at least at a minimum standard.”

Monitoring comes with additional requirements for the laboratory, Dr. Meserve says, and that’s why the CAP has compiled an extensive FAQ resource to accompany the new monitoring status. The first frequently asked question is one of the most pertinent: How laboratories should conduct validation and verification. “The issue there is that if labs are already running these assays in-house, they may want to consider increasing the validation documentation to meet the standards set forth in CAP Center guidelines,” Dr. Meserve explains. “If the lab is not currently performing these assays, then they will have to be prepared to meet those requirements.” And, she notes, large validation cohorts are a big ask in certain situations when a positive result is an uncommon event.

The addition of requirements associated with monitoring status may affect whether laboratories perform an assay in-house. Dr. Meserve’s primary institution, Maine Medical Center, is the largest medical center in Maine, but Maine is a state with a relatively small population. “We do not run highly sensitive ALK in-house here because we have not historically had the testing volume to justify validating this IHC assay. If now, due to monitoring status, the case requirements for validation were to increase, we likely could not meet that requirement locally.” Other laboratories may decide the new monitored status is reason to start using a reference laboratory for these biomarker tests.

Those factors have helped make “When to retire an assay from a test menu” one of the FAQs. “The fun part of IHC for most laboratory directors is validating new assays,” Dr. Meserve says. “It helps keep your lab current, and if there’s a new marker out that can help us make new diagnoses, we want to experiment with it. We want to learn from it. But if an antibody’s diagnostic utility has decreased over time, and other research suggests it is not as specific as hoped when it first came out, then laboratory administrators may suggest retiring it if we order it only a few times a year or if it’s become less useful—or both. It’s just as important to be sensitive to the bandwidth of your laboratory as it is to be bringing on new tests all the time.”

The FAQ document is an attempt to say something helpful but not laboratory-specific about how laboratories performing IHC could approach technical issues, Dr. Meserve says. The list of questions was compiled from Surveys participants’ more challenging questions.

A large number of the other questions address what to do if the laboratory has an unacceptable response on a proficiency test. “Those questions are perhaps the most important of the FAQs,” Dr. Bellizzi says.



Dr. Meserve

An unacceptable response, says Dr. Meserve, “means they got one tumor core wrong out of—almost always—10 tumor cores. So they’re 90 percent correct. Eighty percent is the passing threshold for most assays, except for the monitored ones, which is higher. There are many reasons why a laboratory may have one unacceptable response. And not all of them are going to require changing the assay.”

“In contrast, if a lab were wrong on four cores out of 10, you definitely need to do something,” starting with an investigation, she says. Any concern about the performance of an assay in the laboratory should trigger at minimum an informal process improvement assessment (PIA) to determine the cause and triage the problem

appropriately, the FAQs note.

The FAQ document outlines the “8D” approach to process improvement, similar to the steps included in Lean/Six Sigma process improvement, through which a team can conduct root cause analysis; devise, implement, and validate permanent corrective action; and prevent recurrence. Included in the FAQs are specific steps following a hypothetical “unacceptable” response on proficiency testing for ER, PgR, ALK, BRAF V600E, and KIT when corrective action is needed.

Action may not be needed in every case. But, as a laboratory director, Dr. Meserve says, “My job in the lab, if we do nothing, is to be able to defend a decision to do nothing. I may make a professional judgment that nothing was the appropriate course of action, and I think that’s hard for some people to sit with. They want to do something. In my opinion, the ‘something’ is the investigation and there may not be a necessary corrective action.”

Nevertheless, with the FAQ document, “We’re trying to make the point that if you have even one unacceptable response, you should take a look at the data,” she says. “What sometimes happens is you have one that’s clearly unacceptable but you have three that are in the borderline range and six that are clearly acceptable. And if there is a trend toward being in this unacceptable category, I would argue you need to look into this. This is your bellwether. This is the canary in the coal mine. I think that’s helpful for laboratory directors, and they should investigate that. Otherwise, they might complete the same Survey again and completely fail, because all three of those borderlines will have transferred to the unacceptable category if there’s drift.”

What assays might be next in line to become monitored? In Dr. Meserve’s view, “It would have to be a predictive marker, something used to drive decisions about therapy for patients. So I think it would be HER2 in other organ systems—certain gynecologic malignancies, lung cancer, sometimes colon cancer. Those three organ systems might have their own criteria, different from breast cancer, for determining positivity. Monitoring of that, even though it’s the same analyte in different organ systems, could be very relevant if the assays are being used to decide treatment, and especially when the thresholds for interpreting them as positive or equivocal or negative are different.”

For the time being, laboratories should do a few things to prepare for the new PT requirements, Dr. Meserve says. “Number one is to generally become aware of the new requirements. Number two would be to subscribe to the now-required proficiency tests. The window just opened for subscriptions for next year. So if people want to participate in the first round of highly sensitive ALK lung testing next year, they have to start enrolling now.” Number three would be to pay attention to other guidance documents, in particular the frequently asked questions. “Lab directors like me,” she says, “should be trending rates of positivity in gastric HER2 results as well.”

When the updated guideline “Principles of Analytic Validation of Immunohistochemical Assays” is released next year, “we’ll need to edit the FAQs to reference that document,” Dr. Meserve says. “In some ways the FAQs are kind of an interim product until that happens.”

But the process improvement assessment component of the FAQs is not likely to be addressed in the guideline. So that section of the FAQs will continue to provide important guidance to laboratories in maintaining quality of testing of these biomarkers, she says.

“Laboratories have many other resources to use for quality improvement, and the FAQ document is just another opportunity, an attempt to consolidate the opinions and knowledge base of the committee into an accessible, somewhat informal document.”

That document, and the new requirements for monitoring gastric HER2 and highly sensitive ALK in NSCLC, are each meant, she says, to be a tool in the laboratory director’s quality toolbox.□

Anne Paxton is a writer and attorney in Seattle.