Breast cancer breakthrough sparks HER2 quest

Karen Titus

June 2022—The latest advance in breast cancer treatment is a big one—the promising antibody drug conjugate fam-trastuzumab deruxtecan-nxki, or T-DXd (Enhertu). The drug was granted breakthrough therapy designation this spring for patients with HER2-low metastatic breast cancer, and the drug and trial on which the decision was based were the focus of the plenary session at the ASCO annual meeting in early June.

“This drug in particular is a variant of a drug we are all very familiar with—Herceptin, or trastuzumab,” says David Rimm, MD, PhD, the Anthony N. Brady professor of pathology, professor of medicine (oncology), director of the translational pathology and Yale pathology tissue services, and director of the physician scientist training program in pathology, Department of Pathology, Yale University School of Medicine.

Also familiar: the IHC test to determine eligibility for the drug, a companion diagnostic developed decades ago.

But that’s where easy familiarity ends. Long accustomed to looking for IHC levels in the higher range of expression (3+, 2+) to qualify patients for Herceptin, pathologists might now have to turn their expertise to sorting cases in the low range of protein expression (IHC 0 versus 1+) to identify patients who could benefit from Enhertu (Daiichi Sankyo and AstraZeneca).

It’s complicated for clinicians and patients as well, who may not share pathologists’ appreciation of the subtleties of HER2 testing, says Kenneth Bloom, MD, currently a consultant to Nucleai and OneCell Diagnostics. “It feels like we’ve been quantifying HER2 forever.” Pathologists have been reporting HER2 as 0 and 1+, “giving the impression that we could reliably quantify ‘low’ HER2 results. Our current assays are not really linear in that range at all,” he says.

Peeling back the promise and predictions, pathologists have begun envisioning low-HER2 testing scenarios. Will the current IHC assays suffice? What more can or should be asked of them? Can the assays be retuned? Should they? Will low-end testing require a second test, à la FISH for IHC 2+ cases? Behind every question lie others, usually of the *So the real question we need to answer is* ____ variety. And beneath it all is the acknowledgment that there’s just not enough data right now to decide anything.

The new drug consists of an antibody, a cleavable linker, and a cytotoxic topoisomerase I inhibitor, Dr. Rimm explains. Enhertu doesn’t require HER2 amplification to be effective, but it does require HER2 protein expression. Patients who truly express no HER2 are unlikely to show benefit from the drug. And its toxic side effects and adverse events can be severe.

Current IHC assays are designed to detect amplified versus nonamplified HER2. A nonamplified sample will look like a 0. “We can’t tell the 0’s from the 1’s very well because the assay is not designed to perform in that range,” Dr. Rimm says.

Explaining further, he notes that a 3+ is roughly 5 million molecules/cell; a 2+ would be about half a million. Identifying a 1+ is much harder—“probably 50,000, or maybe 100,000—who knows?” he says. Normal breast is probably 20,000, but there are breast cancers that appear to have no HER2 at all.

Dr. Kimberly Allison (above at Stanford) and other experts in breast pathology and HER2 assessment are now turning their focus to the differentiation of low levels of HER2 protein expression. [Photo by Cindy Charles]

“So what we need is to tell 1,000, or let’s say less than 1,000, from 20,000 or 50,000 or 100,000,” Dr. Rimm says. “And that commercial assay doesn’t exist.”

That puts pathologists in an awkward position, to say the least. The assay used in the clinical trials, Destiny-Breast03 and Destiny-Breast04, is an FDA-approved IHC assay but not one designed for the dynamic range in which it needs to work.

Can pathologists make it work? As Dr. Bloom puts it: *How good do we think pathologists are at reproducibly identifying the 1+ cutoff?*

Dr. Rimm and others set out to find out, in a study published this year in *JAMA Oncology* (Fernandez AI, et al. *JAMA Oncol.* 2022;8[4]:1–4). In the study, Dr. Rimm and his coauthors used CAP proficiency testing data from 2019 and 2020, from more than 1,400 laboratories, to explore how well participants did at distinguishing between 0 versus 1+ cases. They also assessed data from a Yale University study of concordance of 18 pathologists reading 170 breast cancer biopsies.

In the CAP data, 65 percent of the 80 cores evaluated had a concordance rate of 90 percent or greater. Notably, this agreement rate was for scores of 0 and 3+. Of the 80 cores, 56 were considered negative (score of 0 or 1), and in 25 percent of those cores, the agreement was less than 70 percent. In the Yale cohort, disagreement between 0 and 1 was significantly larger than for 2 versus 3.

“When you use the existing assay, with its current dynamic range, pathologists as a group can’t tell 0’s from 1’s very consistently,” says Dr. Rimm of the Yale data.

The CAP data was similar. “It’s pretty much a coin flip in the low cases,” Dr. Rimm says. “In the high cases we do

fine."

Outside the setting of a clinical trial, a 0 and a 1 are considered negative results, and distinguishing between them does not make a difference in treatment, says Kimberly Allison, MD, director of breast pathology, professor of pathology, and vice chair of education, Department of Pathology, Stanford University School of Medicine.

The differences between the two can be subtle, and distinguishing between the two has no bearing on proficiency testing performance.

"That's not a threshold we look at," says Dr. Allison, who coauthored (with Antonio Wolff, MD) an accompanying commentary in the same issue of *JAMA Oncology*. "There's no 'ding' for calling 0 versus 1+. If anything, it's surprising there was decent concordance in that setting."

But with clinical implications evolving, so too are views of the assay and how it might be used.

"The crux of the problem is that the drug is going to get approved with the assay that's not designed for that range," says Dr. Rimm (who, along with the other sources, spoke with CAP TODAY before the drug was approved and before the ASCO meeting this month).

The problem lies neither with the IHC assays nor the pathologists using them. The current assays do what they're supposed to do—asking why they don't do something else is like taking umbrage at a vegan restaurant for not having steak tartare on the menu.

Cautions Dr. Allison: "We don't want to mess up what we're already good at doing—defining 3+ versus not. You don't want to fit your test to two different purposes."

The current fit-for-purpose assays were developed to identify patients who show HER2 overexpression or amplification of the *HER2* gene. "We're actually quite good at doing that," agrees Dr. Bloom. "We've spent a lot of time and effort as CAP, and as pathologists in general, identifying how to perform and interpret those assays. And, of course, they're the FDA-cleared companion diagnostics for that purpose. CAP data shows we do a good job of reproducibly identifying cases that show overexpression of HER2, meaning they express at the 3+ level."

In practice, Dr. Bloom notes, pathologists tend to review the HER2 slide at relatively low magnification. Tumors that show strong overexpression are easily visible—pathologists can see the so-called chicken-wire pattern and the complete, circumferential membrane staining of HER2 on the tumor cells, which helps identify tumors that are truly 3+ versus a pattern that mimics 3+ expression but is not truly amplified. When the complete membrane staining is only seen on higher magnification, the tumors are categorized as 2+ and reflexed to FISH.

In clinical practice, there is little need—until perhaps now—to go to a higher magnification. In the clinical trial for Enhertu, says Dr. Bloom, 40× was used to distinguish 0 and 1+ expression in low-expressing cases.

Rather than providing an easy solution, however, adjusting the magnification gives rise to another layer of questions. "How low of a HER2 expression do you need to see to show efficacy of this new class of antibody conjugate therapies?" Dr. Bloom asks.

That won't be easy to answer. Antibody drug conjugates have their own peculiarities. In the case of Enhertu, when the drug identifies a cell expressing HER2, it becomes internalized within the malignant cell, Dr. Bloom explains. The topoisomerase I inhibitor traverses the cell's membrane after it kills the cell, creating a so-called bystander effect that kills cells in the surrounding area, even if they don't express HER2.

And now, the deluge.

Though the mechanism is not fully understood, and the data nascent, "there's this concept of the importance of the spatial relationship of HER2-expressing cells to other tumor cells," Dr. Bloom says.

"When we use a criterion, such as expression in at least 10 percent of tumor cells," he continues, "the relationship

between HER2-expressing and non-expressing tumor cells would be wildly different if the 10 percent of expressing tumor cells were all clustered in only one tiny area, versus being randomly scattered throughout the tumor." The data so far suggests the drug's efficacy appears to be related to the expression level of HER2 throughout the tumor, but this is hardly settled business.

The stakes are high. One concern is that patients in the metastatic setting might not receive a drug that could help them. In the adjuvant setting the concern is the reverse: They may receive a targeted therapy when they do not express the target, and thus not benefit from the treatment. Since the trial was for the metastatic setting, Dr. Rimm notes, the first concern might be more likely, "but it's likely both events are going to happen," he says, though preliminary data suggests undertreatment will be the main problem. At CAP TODAY press time, the drug was expected to be approved soon for patients with unresectable or metastatic HER2-low breast cancer who have received prior systemic therapy in the metastatic setting. In the future there may also be indications in the neoadjuvant or adjuvant setting for patients who have disease recurrence during or within six months of completing therapy. (The drug was approved in May for patients with unresectable or metastatic HER2-positive cancer who received a prior anti-HER2-based regimen.)

That puts plenty of pressure on pathologists to sort matters out. "This was a relatively hot topic at the USCAP annual meeting earlier this year," says Dr. Allison. "Pharma had a whole session on trying to get us to score accurately between 0 and 1+. That drew a lot of comments."

"There's a lot of hype right now about us as pathologists being able to reproducibly differentiate 0's versus 1+'s," she adds. "I'm hoping that will be an irrelevant threshold. But maybe it won't. It's hard to predict."

If it is relevant, she continues, pathologists can learn to be reproducible. "But I worry more about subtle issues that may cause a 0 versus a 1+ result," including the preanalytic issues that can affect IHC, particularly in the lower range. That includes ischemic time, antibodies, and whether the sample comes from a core biopsy versus the excision of the primary, or whether the metastatic site was used for the sample.

Down the road, she suggests, pathologists may face questions about what to test. "Can we test the primary if the metastatic sample is not available?" She suspects the answer will be yes, because that has been true in other trials. "But I wonder which is most likely to have 1+ staining." Core biopsies might have the highest likelihood, she says, because they have the shortest ischemic time. She says she's talking to her colleagues about all these matters.

It's difficult to know whether pathologists in the study published in *JAMA Oncology* would have done better at distinguishing 0's and 1's had they known to look more closely at those cases. Providing that instruction would have biased the study, Dr. Rimm says. "There really wasn't a good way to handle that weakness. We didn't want to try to make the pathologists pick apart that dynamic range, since that's not what the assay was designed for."

Moreover, Dr. Rimm says, focusing on how well pathologists could do in that range "would be almost meaningless, because if they could get some consistency, we don't know if they would be picking the responders or the nonresponders," given that they weren't looking at specimens from the clinical trial. That information would be better coming from AstraZeneca and Daiichi Sankyo, he says.

The Destiny trial did not appear to include patients whose tumors were classified as 0, according to those interviewed for this story. But another study, known as Daisy, did, and since some of those patients with scores of 0 did respond, "that makes us think those were really 1+'s, and that's the concern," Dr. Rimm says.

Echoing Dr. Allison, Dr. Bloom suggests that pathologists could become more reproducible than the study published in *JAMA Oncology* suggests. But that's not the point. The current assay is simply inappropriate for newer needs, he says.

"It's not an antibody problem. It's basically a titer problem," Dr. Bloom says. "We need a much more concentrated dilution so that we can be linear in the low range of HER2 expression." But doing so would skew the number of

cases reported as 3+. “We would start calling tumors 3+ that aren’t really overexpressed because the assay would saturate much too quickly.”

It seems unlikely that pathologists can reliably quantify both high and low ranges with the same assay, he says. Burning both ends of the candle may seem like a bright idea, but it’s not a long-term strategy.

Nor is trying to answer questions with insufficient data.

“What’s the best way to measure HER2 low?” Dr. Bloom asks. For that matter, what *is* HER2 low? “We’ve only had glimpses of data on HER2 low from the clinical trials.” Within that limited data, it appears that the higher the HER2 expression level, the more effective the therapy seems to be—including patients who might typically be called negative by IHC.

On a practical level, pathologists are trying to detect HER2 molecules on cells in the range from up to a thousand to several million, as Dr. Rimm has noted. Adds Dr. Bloom: “That’s a 3-log difference. That’s more than a thousand-fold difference between low-expressing and potentially high-expressing tumors.” Current IHC assays that rely on chromogenic detection systems have the ability to detect about a 1–1.5-log difference in expression, he says, meaning they can detect about a 10–50-fold difference.

Dr. Bloom and Dr. Rimm are coauthors of a paper that demonstrates that changing the level of antibody concentration may enable the current assays to be linear at low levels of expression instead of at high levels (Moutafi M, et al. *Lab Invest.* Published online ahead of print May 20, 2022. doi:10.1038/s41374-022-00804-9). Says Dr. Bloom, “You still have that 1–1.5-log expression level difference, but you have to decide: Do you want to be linear in the lower range of expression, or do you want to be linear in the higher range?”

That’s not a theoretical issue. Will pathologists need to distinguish the difference between 5,000 and 10,000 molecules? The clinical data aren’t available to answer that question, but the trend seems to be that tumors that show higher levels of HER2 expression derive more benefit, Dr. Bloom reiterates.

All of which puts pathologists in their current position, says Dr. Bloom: “When we call HER2 0 with our current assay, which we know isn’t linear in that lower range, it means that some of the cells we’re calling 0 might have significant HER2 expression.” Likewise, he says, calling a tumor HER2 1+ means it’s likely that the tumor has higher expression levels than a tumor assessed as HER2 0. In Dr. Bloom’s opinion, most cases that are called 1+ are at least 1+. But a 0 may not in fact be a 0. Says Dr. Rimm, “This suspicion is confirmed in the paper in *Laboratory Investigation.*”

What truly is 0 versus 1+? Dr. Allison teases out the question further. “Should 0 be below a normal breast epithelium? Because normal breast epithelium has some HER2. You’ve got to come up with some standard to calibrate to if you’re going to make that a threshold.

“But I wouldn’t want to recalibrate all our assays, which are so fine-tuned to predict HER2 positive versus not,” she continues. “We might need a separate HER2 test calibrated for the metastatic setting. Or maybe we won’t.”

So what’s a pathologist to do?

Using the current assay at 40× “sounds interesting, but there’s no evidence to support it,” Dr. Rimm says. “I guess you have to decide whether you’re depending on evidence or whether you’re depending on what sounds like a good idea.”

Other options are under discussion. Dr. Rimm suggests it would be reasonable for vendors to retune their assays, adding more antibody to improve sensitivity. Once a bridging study is done, that could provide a solution, he says.

Dr. Bloom says it makes sense to continue to use the HER2 FDA-cleared antibodies for assessing cases for high levels of HER2.

If those cases are not called positive, then pathologists will need to decide whether to run a second assay to

determine HER2 low levels—and whether those results will be useful. Does it make a difference if there are 60,000 receptors or 40,000, or 20,000 versus 10,000?

Dr. Bloom

Regardless of the eventual answer, Dr. Bloom says, “You won’t be able to see that difference with the current assays because they are not linear in that range. It doesn’t make a difference what you do. It doesn’t make a difference what games you try to play. It’s fundamentally the wrong assay.” On a practical level, “That means when we call something 1+, the range of HER2 expression in that 1+ could vary significantly. Meaning, some of the 1+’s may have as many as 200,000 receptors, while other 1+’s only 50,000 receptors.

“That’s a pretty big difference,” he says. “Yet for an oncologist reading a report, the results will be the same.” But if the number of receptors is what’s determining the efficacy of the drug, a fourfold difference in expression should be important.

Pick the range that addresses the clinical question, Dr. Bloom says—once you know what that range is. Which no one currently does. “Right now we just have trends looking at low levels of HER2 expression with nonoptimized assays,” he concedes. But if the data moves forward, the current assays run the risk of missing patients who would be eligible for Enhertu. “We could be missing a large number of patients and potentially mislead oncologists by including a pretty wide expression level of HER2 in the 1+ category.”

Continuing the journey in the Land of Ifs, Dr. Bloom returns to the matter of bystander effect. If that is the working mechanism, the question becomes: What is the arrangement of cells showing expression within the tumor? What is the distribution? Early data suggests this could have potential clinical implications. But for now, Dr. Bloom says, there are only abstracts—without sufficient data to answer questions about which assays might be appropriate. “They’re just teasing us right now.”

Dr. Rimm talks about using digital image analysis to measure HER2, which the CAP’s data indicates about 20 percent of laboratories already do. Digital assessment might function almost like FISH does currently for borderline 2+ cases of HER2 IHC. “A 0 would be like a 2+. That’s what we need—some way to adjudicate the 0’s. It would be a second step. It’s just that the second step hasn’t been proven yet, although the *Laboratory Investigation* paper describes a fully quantitative low-range assay that could serve this purpose.”

There are many advantages to image analysis, Dr. Allison says, but they’re matched by the need for caution. Just as it’s possible to increase the intensity of stains for IHC, “the same can be true for intensity of staining using digital image analysis.” Setting the thresholds is still key. “So there are pitfalls. It’s imperfect in different ways.” It also adds complexity and time, but it might be useful as a second read, or as quality control, she says.

These and other issues will likely wax and wane as antibody drug conjugates become part of the scenery, but for now excitement lingers in the air.

Dr. Rimm predicts Enhertu will change the landscape of treatment for breast cancer once it gets into the adjuvant setting. Fifteen percent of patients qualify for trastuzumab; as many as 80 to 90 percent could qualify for trastuzumab deruxtecan. He reports his oncology colleagues are calling the drug “the biggest thing to come along in breast cancer since chemotherapy.”

Even apart from the specifics of the Enhertu story, pathologists need to prepare to become quantitative with their measurements, Dr. Rimm insists, and not simply rely on ability to read cases as 0, 1, 2, and 3. He says he prefers a

concentration on a tissue, versus counting molecules per cell. With the latter, he says, “You don’t know whether you’re cutting through cells at the North Pole or the Equator.” Averaging cells across a defined area, such as per square millimeter, on the other hand, is a “quantitative assay you can kind of hang your hat on.”

Dr. Rimm

That’s what pathology lacks right now, he says. “As pathologists, if we want to keep up with the times, and we want to keep getting our business from our oncologists, they need this kind of accuracy with this new class of drugs. We can’t rely on just our eyes being able to judge intensity like we did in the last century,” he says. Pathologists have long made it their business to “judge stuff—that’s how we make our living. But if we keep judging stuff, it will be taken out of our hands. They will take specimens somewhere else to get them measured.” Digitizing and quantification are not that hard anymore, he says. “If pathologists won’t do it, someone will.”

True, says Dr. Bloom. The world of spatial transcriptomics and spatial multiomics is garnering interest with the advent of new immunotherapy combinations and antibody drug conjugates. “It’s not prime time—yet,” he says. “But at some point it’s going to be incumbent on us to use these tools. The emergence of spatial biology is real. Pharma is paying a lot of interest into how spatial biology affects the efficacy of their therapies. We can likely expect, as pathologists, that in three to five years that will translate to us.”

And in the meantime? Even on the more familiar territory of HER2 assays, says Dr. Bloom, “Until we see more data,” what’s actually needed “is just a guess. We’re getting ahead of ourselves until we see the data.”

But the current discussion is important, he says. “This is exciting. Breast cancers showing low levels of HER2 expression are seen in a significant number of women. Over 80 percent of breast cancers are going to be potentially eligible for this therapy.

“There’s a lot of noise around this,” he concludes. “And there should be. This is a big deal.”

Karen Titus is CAP TODAY *contributing editor and co-managing editor.*