Low level limbo in HER2 breast cancer

Karen Titus

August 2023—Seemingly channeling the inspiration of Magritte and his famous pipe, pathologists are painting a new picture of what has long been an everyday object in their own world: HER2.

To paraphrase the master: *Ceci n'est pas facile.*

For years, HER2 testing in breast cancer has seemed self-evident, ever since the HER2-targeted therapy trastuzumab and its companion diagnostic arrived on the scene a quarter of a century ago. Pathologists became comfortable using immunohistochemistry to identify 3+ cases and turning to in situ hybridization techniques to sort through less obvious ones.

But early last summer, a variant of the drug, trastuzumab-deruxtecan (T-DXd), shook up that routine. When researchers presented results from the Destiny-Breast04 study at the 2022 ASCO annual meeting, showing that T-DXd significantly improves survival in so-called HER2-low metastatic breast cancer, attendees responded with a minutes-long standing ovation.

They then returned from the meeting like evangelicals from the revival tent. "The quest started then," says Shabnam Jaffer, MD, chair, Department of Pathology and Laboratory Medicine, Lenox Hill Hospital/Northwell Health, New York City, and professor of pathology, Zucker School of Medicine, Hofstra University. A case that was 1+ IHC or 2+ IHC/ISH-negative meant patients qualified for the antibody-drug conjugate. "So the oncologists were all very excited and motivated to start labeling which patients were HER2 low. They wanted to treat those patients, especially those who had failed other treatments."

The motivation was stronger than the methods, as pathologists were quick to realize. It was no longer appropriate to lump the 0 and 1+ cases together as "negative," but the assays that had worked well for identifying cases that were strongly HER2 positive were never meant to parse the particulars at lower levels.

ASCO and the CAP moved quickly as well, and this spring the steering committee published an update to the ASCO-CAP guideline on HER2 testing in breast cancer (Wolff AC, et al. *Arch Pathol Lab Med.* Published online June 7, 2023. doi:10.5858/arpa.2023-0950-SA).

"It was much anticipated, much awaited by pathologists," says Dr. Jaffer, "because we were starting to get bombarded by requests and questions from our oncologists, asking about both current and historic specimens: *By the way, that patient you signed out is metastatic—can you go back and let me know if it was HER2 low or not?* Many institutions were in a flurry to figure it out."

Prior to the Destiny-B04 trial, there was no incentive to provide granularity to these reports, she notes. Current methods didn't make it easy to separate 0 and 1+ cases, nor was there a need to; it made sense for testing results to be perceived as binary. "But now that philosophy is changing as we see these new drugs work in patients who are not necessarily HER2 expressed or amplified," Dr. Jaffer says.

The update, the authors make clear, is not a revision. The guideline remains in force. It does, however, offer five best practices to help pathologists and their clinical colleagues navigate the more immediate HER2 testing challenges raised by antibody-drug conjugates.

But these are only first steps in a field that everyone agrees could look completely different in the not-too-distant future, as pathologists decide what to look for, and how.

Dr. Shabnam Jaffer at Lenox Hill Hospital. The recently published guideline update on HER2 testing in breast cancer was "much anticipated, much awaited by pathologists," she says, because oncologists have had many requests and questions about current and historic specimens. [Photo by Jennifer Altman]

There's also the question of what to call these cases. HER2 low is not a clinical category, though the term has become common parlance among physicians. With new ways to assess HER2 in development, as well as dozens of antibody-drug conjugates ready to pop out of the pipeline, "HER2 low" may be a placeholder, eventually giving pathologists the Adam-like responsibility of naming what lies before them.

Despite the locomotive rate at which this all seems to be evolving, the guideline is notable for what it doesn't change, says Kimberly Allison, MD, an author on the update.

"We opted not to create a new reporting category of HER2 low," says Dr. Allison, director of breast pathology, professor of pathology, and vice chair of education, Department of Pathology, Stanford University School of Medicine. Instead, the authors wanted to call attention to the difference it makes clinically when a case is scored as 0 versus 1+. They suggest pathologists use a report comment, or footnote, with proposed wording. "That was our way of explaining that patients who don't have HER2 overexpression or amplification may be eligible for these other therapies," she says.

"It's pragmatic guidance to pathologists in terms of how to deal with HER2-low breast cancers and their categorization," says Stuart Schnitt, MD, chief of breast oncologic pathology, Dana-Farber Brigham Cancer Center; associate director, Dana-Farber Cancer Institute-Brigham and Women's Hospital Breast Oncology Program; and professor of pathology, Harvard Medical School. "There's nothing radical that we need to do at this point," but rather, continue to follow the 2018 ASCO-CAP guidelines, while noting the more focused emphasis on figuring out the difference between 0 and 1+ cases. (Dr. Schnitt is coauthor of an editorial published in the same issue as the

update: Schnitt SJ, et al. *Arch Pathol Lab Med.* Published online June 7, 2023. doi:10.5858/arpa.2023-0187-ED.)

As Dr. Allison walks through the five best practices to help pathologists make that distinction, their workaday nature quickly becomes obvious.

No. 1: "Basically you want to use ASCO-CAP scoring criteria for IHC HER2," which is what researchers used in the Destiny-B04 clinical trial. "There's no new scoring criteria or algorithm."

But since the criteria were not designed to qualify patients for antibody-drug conjugates, that brings up best practice **No. 2:** Pathologists should examine HER2 IHC at high power—40×—when trying to discriminate 0 versus 1+ staining. That's now a clinically relevant, although not validated, threshold for whether to treat metastatic patients who are eligible for T-DXd. For now, Dr. Allison emphasizes, the focus is on patients with metastatic disease (though that could change). "But we want to be diligent for all of our scoring for all samples, including primaries, because that data may become relevant if the patient becomes metastatic."

No. 3 suggests that for cases that present a scoring challenge, consider having a second pathologist review cases that are close to the interpretive threshold, i.e. greater than 10 percent of cells with incomplete membrane staining that is faint/barely perceptible.

Another trial, Destiny-B06, is looking more closely at those hard-to-distinguish cases, Dr. Allison notes. "They're terming cases with some partial membrane staining but that falls just below the 10 percent threshold for 1+ 'ultralow.' But that's for the future." For now, she says, "We just do our best using the criteria we have." Having a second pathologist is similar to recommendations that call for a second pathologist to review borderline ER cases. "Anytime you're close to a threshold it's a good idea to consider having someone else agree with you or disagree with you, so you know you're all resulting similarly in your practice," Dr. Allison says.

Best practice **No. 4** calls for using controls with a wide range of protein expression, including 1+, to ensure an appropriate level of detection. Dr. Allison says on-slide controls are recommended, though not required.

At Stanford, she and her colleagues use a tissue microarray slide control, with multiple dots of tissue with different IHC expression including 0 and 1+. "It's a little bit of a moving target," she says, "because the clinical trial didn't really test the threshold between 0 and 1+ so there are no clinically validated controls to calibrate to. But you do want to make sure your assay is not only able to pick up the highest levels of 3+ protein expression, so including cases that test 1+ in your lab will help ensure your assay is not becoming less sensitive."

No. 5 guides labs to look at the preanalytic conditions to ensure that samples from both primary and metastatic sites have appropriate fixation and cold ischemic times. That can be a particular problem for metastatic sites, Dr. Allison cautions, since when a lung lesion, for example, or brain or liver metastases are biopsied, the pathologist may not know up front that the primary tumor is breast cancer. "If breast cancer is suspected in any metastatic site, pay attention to those preanalytic variables, since they can affect HER2 testing, and probably other tests as well," she says.

How difficult will it be to put these best practices to work? "It depends on what's within the pathologist's control," Dr. Jaffer says. "I can control the magnification I'm going to use to inspect the slide." Preanalytical details are another matter. "Setting up standards for the multidisciplinary team is initially challenging. But with due course of time and action, it gets better."

Using 40× to look at 0 and 1+ cases should be "fairly standard," Dr. Allison says. "It's just a matter of how hard you are looking"—like the difference between witnessing a bank robbery and tracking down embezzlement. "Before, we were looking for the high end of the spectrum, which was obvious. You could see it—it was very strong." At the lower end, "you kind of hemmed and hawed," she says.

Second pathologist reviews are fairly standard as well, though it doesn't guarantee consensus.

Studies have shown that agreement is "pretty poor if you're unaware there's a difference between 0 and 1+," Dr.

Allison says. “But when you pay attention, and there’s more training, it’s higher.” Nevertheless, “Borderline cases are borderline cases,” she says, and are likely to have pathologist-to-pathologist disagreement.

So the point, Dr. Allison continues, is, “While we may not be sure what the most accurate result is, make sure you have concordance at least in your group.” Otherwise, the oboe has left the stage: “There’s nothing else we can fine-tune to.”

Ideally, Dr. Jaffer says, the second pathologist will have expertise in breast—“someone who does it day in, day out.”

Dr. Schnitt recalls discussing borderline cases with his colleagues during his work as a member of the pathology group for the European Society for Medical Oncology guideline. “Say you have six pathologists looking at the same case, and half would call it 0 and half would call it 1+. Do you err on the side of overcalling, or do you err on the side of undercalling?” he asks.

Dr. Schnitt

He suspects pathologists tend to overcall cases that are on the fence. “I talked to an oncologist from Italy several months ago, and he told me he’s basically not seeing any HER2 0s anymore,” he says with a laugh. “People are afraid they’re going to miss something that’s going to benefit the patient, and none of us wants to do that.”

In his own practice, Dr. Schnitt has turned to his oncologists for guidance. “I’ve specifically asked them, in borderline cases, would they rather have us undercall or overcall?” Most lean toward overcalling, he says, though “they’re very cagey about it. They basically say, *We want you to be accurate; we want you to get it right.* But the reality is, in these borderline cases it’s difficult.” And while medical oncologists are quite savvy, he says, “Some of them don’t understand that we’re using a test to make a distinction that wasn’t designed for that purpose.”

Using a wide range of controls might be the hardest of the best practices to, well, practice, Dr. Schnitt says. “The toughest thing may be to get controls that are 1+. Not that it’s terribly difficult, but people have to change the way they do controls.” Options include using cell lines that are embedded in cell blocks, he says, or using tissue sections embedded in paraffin that are 1+.

As for adhering to preanalytic conditions, Dr. Jaffer is a stickler. This is routine with primary tumors, she notes, but now, “if we’re getting a sample from a metastatic site, we need to let the oncologists and the person who’s procuring the tissue know that the same principles apply.”

She pursues a low-tech approach to convincing her clinical colleagues, using pictures from the literature to show the impact of delaying ischemic time. “When you show that to the surgeon, pictures are louder than words. They get it. And they start working with you to achieve it.”

Dr. Allison

The call for a footnote acknowledges the linguistic pachyderm in the room: As the authors note, “It is premature to change reporting terminology for lower levels of HER2 IHC expression,” even as scrutiny of these cases has increased. The goal was to avoid the confusion of introducing terminology that is likely to continue to evolve, and which hasn’t been validated.

Medical oncologists know they have new treatment options and will understand the implications of the various IHC scores, Dr. Allison says. They’re not necessarily looking for the words “HER2 low” and can get the information they need from the IHC raw score.

HER2 low may be a convenient and easy-to-say label, but it’s not a new subtype of breast cancer. “I think there’s a lot of misinformation about that—that there’s a new subtype, and we have to identify it,” Dr. Allison says.

But the term has become commonplace and, as noted, was used in Destiny-B04. Pulling it from use seems akin to pushing a barrel back up to the top of Niagara Falls. The irony is not lost on Dr. Allison. “The goal of the update was to end confusion,” she says with a laugh. “But the phrase is out there. And let’s be honest—it’s a great marketing tool, right? *There’s a new subtype, and we need to find it.*

“We agree with the need to pay attention,” she continues. “But we need to answer more questions before we use new terms. For example, we really don’t yet know if some or all IHC 0 cases are also ‘HER2 low’ since they were excluded from the trial.”

By now, most oncologists already know the clinical relevance of reporting specific HER2 protein levels, Dr. Jaffer says, and know the criteria for the lower levels. But in her experience, they’d also prefer a shorthand answer, “like we do for the HER2 positives,” she explains. “It would help for them to have a quick identifier in the report that states what is HER2 low, even though they can glean that information from the reports.”

She understands the limitations on the update’s authors, including the fact that it’s not clear yet how HER2 0s behave. “But the labeling of HER2 low is so entrenched in our literature already,” says Dr. Jaffer, who wrote about her concerns in an editorial (Jaffer S. *Arch Pathol Lab Med.* Published online June 7, 2023. doi:10.5858/arpa.2023-0176-ED). “We discuss it as a terminology in our multidisciplinary breast discussions and at conferences. At the recent USCAP meeting it was all over the place. Everybody was using that lingo.”

It’s a waiting game to see how the unofficial but widespread label will evolve. “As more drugs become available with the same payload effect, we’ll see more and more patients responding,” Dr. Jaffer predicts. And if that includes HER2 0 patients, “then maybe we don’t have to get into this murky territory.” The ongoing Daisy trial and Destiny-B06 should provide crucial information, including whether HER2 0s and ultralows respond, “and whether it’s important to even have this HER2-low labeling or not,” Dr. Jaffer says.

“HER2 low” isn’t the only beleaguered bit of vocabulary. Dr. Jaffer isn’t alone when she says, “I’ve always had this problem: the fact that we lump 0 and 1+ together as negative. I think the time has come to separate those out and call the 0s something else—null as currently used or dead negative, or something of the sort. Because the 1+ appears to be expressing something. That is my pet peeve. I wish we could call the negatives truly negative, and then give 1+ a new label.”

As language struggles to keep up with the field, pathologists continue to float new approaches.

In his European Society for Medical Oncology work, Dr. Schnitt says, “There were a number of pathologists in Europe [who] were absolutely adamant about using HER2 low in the pathology reports. They basically thought it was negligent not to,” although the ESMO guideline itself does not use it, he says.

He also recalls speaking with a group of pathologists who argued that the definition of 1+ staining should be changed from “faint/barely perceptible” to “faint,” with the notion that this would improve observer reproducibility.

His response? “It is beyond my comprehension how such a change would improve matters, particularly since the dictionary definition of ‘faint’ is ‘barely perceptible.’”

Given the growing spectrum of drugs that target HER2, the demands put on the classifications will increase as well. The terms positive, negative, and equivocal were developed to identify patients who are likely or unlikely to respond to conventional HER2-targeted therapy, such as trastuzumab or pertuzumab. These same cases may not be negative in terms of response to antibody-drug conjugates. “It seems almost paradoxical to say the case is 1+ and negative,” says Dr. Schnitt. “In some ways it’s even misleading these days to say 1+ is negative. You’re saying it’s negative for the likelihood of response to one kind of drug, when in fact it’s positive for the likelihood of response to another type of drug.”

Echoing Dr. Jaffer, he says he understands the constraints on the steering committee, but adds, “I personally would have been much happier if they said continue to score as 0, 1+, 2+, 3+, but drop the positive, negative, and equivocal.”

Dr. Schnitt also recognizes why guideline groups are hesitant to adopt HER2-low terminology, noting that in current practice, it would be hard to use it consistently if it were based on IHC alone. If a case is 1+ by IHC, the HER2-low label is easy to apply. If a case is 2+, it’s impossible to say whether it’s HER2 positive or HER2 low until ISH results are available. It’s possible a lab could hold on to those HER2 results until both tests are done, he says, “but I don’t know of any institution in the U.S. that does that.”

Dr. Schnitt draws an analogy to triple-negative breast cancer (for which there may also be implications as HER2-negative status evolves). “We never use the term ‘triple negative’ in our pathology report,” he says. Rather, the ER/PR/HER2 negative report leads to the clinical interpretation of triple negative. “I think you could argue that HER2 low is not a pathologic diagnosis—it’s a clinical interpretation of HER2 assay results,” one with implications for therapy.

The lower levels of HER2 suffer from an inconsistent identity, an Alsace sliding between France and Germany. Cases alternate between 0 and 1+ biologically in 30 to 40 percent of cases. “It’s pretty much equal movement in both directions,” says Dr. Schnitt. “One biopsy on a patient can be 0 and another biopsy 1+, either synchronously or metachronously.”

“We’re spending a lot of time thinking about this now,” Dr. Schnitt says. If the Destiny-B06 trial shows a response in HER2 0 cases that are ultralow, “then the only cases that are not HER2 3+ or 2+/amplified that we’d have to identify are the ones that are totally negative, with no staining at all.” To his knowledge, however, there are no current trials that include HER2 0s with no staining, often termed “HER2 null.”

“We may eventually come full circle,” Dr. Schnitt muses. After Destiny Breast-06, it’s possible there will be four categories: HER2 positive, HER2 negative, HER2 low, and HER2 ultralow, he says. But if HER2 null/0 cases ultimately show a response, as noted, the binary approach that evolved in the early days of Hercep-Test/trastuzumab may return. “I don’t know where we’re going to be in two or three years with this.”

Some clinical data suggest that HER2-positive tumors that are heterogeneous are less responsive to HER2-targeted therapy than ones that are not heterogeneous, Dr. Schnitt says. Thus, there may be a reason to refine HER2 testing that is independent of the whole issue of HER2 low, he says, pointing to the work of David Rimm, MD, PhD, and others. “Many AI companies are working on quantification of HER2 on a cell-by-cell level, identifying the percentage of cells that show circumferential strong staining, circumferential moderate staining, circumferential weak staining, partial strong, partial moderate, partial weak. And they can give you a histogram of the proportion of cells in each of those categories.”

Dr. Schnitt has heard some pathologists suggest the solution is simply to find an assay that will detect as many HER2-low cases as possible. It’s not a step he necessarily recommends, in part because of the aforementioned reason—who knows who will respond?—but also because it’s impractical. “I don’t think any laboratory is going to change its immunostaining platform because of a single antibody that needs to be used.”

Dr. Jaffer is equally keen to see new technologies improve things. Artificial intelligence will likely play a huge role, she predicts. Interobserver variability can be improved with education, and has been, she says, with options such

as the USCAP tutorials offered on its website. “But human error will never be eliminated.”

The overarching limitations of IHC go beyond problems like interobserver variability, Dr. Jaffer continues. “There may be some low levels that we cannot measure with IHC” but are important for patients.

“We need new assays that stratify cases with lower HER2 expression,” Dr. Jaffer says. These might include methods detecting mRNA, chromogenic and fluorescence techniques, targeted mass spectrometry, immunofluorescence, etc. “All these are new ways to detect HER2 beyond the microscopic level” and could appear in the not-too-distant future. “That’s where I see our push coming—to be able to measure at a level that the pathologist’s eye cannot see.”

As far as Dr. Rimm is concerned, Dr. Jaffer presents the verbal equivalent of a golden ticket: “measure.”

Dr. Rimm, the Anthony N. Brady professor of pathology and professor of medicine (medical oncology), Yale University, has made the future his calling card. “HER2 low is actually a biological category. There’s data for this from both the RNA field and the quantitative protein field to show” there are patients who have no HER2; there are patients who have HER2 present but not amplified; and patients who have HER2 amplified. “And they have a lot,” adds Dr. Rimm, who is also the director of the quantitative diagnostics in anatomic pathology laboratory, the Yale Pathology Tissue Services, and the physician scientist training program in pathology.

Dr. Rimm

That translates to three biological categories, Dr. Rimm continues: HER2 null, HER2 low, and HER2 amplified, none of which lines up neatly with the current scoring system of IHC 0, 1+, 2+, and 3+. In his view, “HER2 low is confusing, but I think it’s here to stay because it’s a biological category.”

Dr. Rimm isn’t shy about how he’d like to see the field evolve. “I’m going to have a tattoo on my forehead that says *Don’t read—measure*,” he jokes. With new methods that have been developed in recent years, it’s become possible to measure protein on the slide using fully quantitative methods.

He and his colleagues at Yale have developed a high-sensitivity HER2 test. They take a signal in the pathologist-circled regions of the slide, then average the signal intensity across the area of the tumor (as defined by cytokeratin expression). This approach generates an assay that is standardized as a measurement in attomoles per square millimeter. “We call this the HS-HER2 assay since it is about 10 times more sensitive than the legacy assays, which puts us in the sweet spot for low HER2,” Dr. Rimm says. With the sophisticated image analysis tools now available in labs, “we can actually measure the amount of protein on a slide using a standard curve, just like we do for a biochemistry assay.”

In the future, he and his colleagues plan to present their work on a duplex assay they’re developing that quantitatively measures HER2 and Trop-2, which is the target of another antibody-drug conjugate, sacituzumab govitecan. With more than two dozen targeted therapy antibody-drug conjugate trials now underway, Dr. Rimm says, the need to measure targets will only grow in the future. Even for drugs that may not require measurement (some say it may not be necessary for Trop-2, he notes), “in fact, if you do measure the target, you pick the patients better than if you don’t measure the target. It’s important for the patients but less emphasized by drug companies because they want to give the drug to everybody.”

“Can you really call it targeted therapy if you don’t assess the target?” he asks.

The duplex test requires fluorescence, but Dr. Rimm has his eye on another prize. His group has submitted an abstract at the San Antonio Breast Cancer Symposium showing how to measure using chromogens. “Pathologists love chromogens because they can read chromogens,” he says. Measuring with chromogens is less scientifically rigorous than fluorescence, he says. “But it’s doable—you just can’t quantitatively multiplex it.”

The HS-HER2 test is available for clinical use at Yale as a validated lab-developed test. “We haven’t advertised it. But we had three patients who found me somehow, so we offered it to them,” he reports. “We don’t know how well it will predict response to therapy—all we know is we can measure it and show protein above the limit of quantification [LOQ] for the analytic assay.”

Obviously many steps lay ahead, and, as Dr. Rimm notes, “It’s quite technical and challenging. It’s not going to be like IHC, where you can just put it on a stainer and take it off and read it. That’s why reading is so popular—because it’s so easy.”

Nevertheless, he predicts measurement will be a reality in the future. The more important question, for him, is who will be doing the measuring? “My concern is that if pathologists don’t adopt the idea of measuring rather than reading, some company will come along and eat pathologists’ lunch.”

Hence the focus on chromogens. “Pathologists are more comfortable with that,” Dr. Rimm says, and it would be easy to integrate into most labs. He also anticipates that it would be easy to incorporate chromogens into the routine digital workflow—another future reality for labs, he predicts.

Though it’s too long for another forehead tattoo, Dr. Rimm offers another guiding precept: “When pathologists think about what tests they’re doing, they need to think about where the field is going.”

There are concerns about offering a drug to those who won’t benefit, though with current testing methods it may not be possible “to keep eating away in the 0 category and finding more patients who will benefit,” Dr. Allison says. At some point there may be a secondary test, a more sensitive method that will enable pathologists to distinguish a low level from, as she puts it, “a zero-zero-zero.”

Dr. Rimm suspects cutoffs will remain important, noting that higher levels of HER2 correspond to greater response rates. In HER2-amplified patients, “just about everyone responds.” In HER2 2+ and 1+ cases, response rates go from 80 or 70 percent to 50 or 40 percent. And in 0 and 1+ cases, the response rate is about 30 percent. “That tells me there probably is some low threshold below which people won’t respond,” he says. “It doesn’t guarantee that every patient above the threshold will respond, but if you have no target, I don’t think this drug works. The bottom line is, there are probably very few patients with no target, maybe five to eight percent,” he estimates.

For now, T-DXd is limited to use in the metastatic setting. Dr. Rimm foresees a future—“not so far away”—when there will be indications for the drug in the early breast cancer setting, which might mean testing the pre-neoadjuvant biopsy. That could open up another option for measuring HER2. “The real importance of a quantitative test will come when we have patients who might be cured by surgery alone,” he says. “It will be important to find that lower threshold below which they have no chance of benefiting.”

The drug could also be used in other types of cancer as well, says Dr. Rimm, pointing to promising work in gastric, endometrial, and cervical cancers. “Already there’s evidence that low HER2 is a phenomenon outside breast cancer. All the cancers that are epithelioid in nature probably express some level of HER2.” Using the HS-HER2 assay, he and his colleagues found naturally occurring HER2 expression in the pseudostratified columnar airway epithelium of trachea. “This may explain some of the toxicity of T-DXd,” he says.

With a different way of looking, it might be possible to see HER2 in unexpected places. As methods and language try to keep up, the field seems to be undergoing a Surrealist moment. HER2 is definitely not a pipe. But, as Dr. Schnitt puts it, “HER2 is not what you think it is.”

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