TUMOR MARKERS

A NEW STANDARD OF BREAST CANCER CARE

Diagnostic test creation often requires using many of the principles of drug development. Complex data analysis and evidence-based reviews are changing professional society guidelines and treatment options for breast cancer.

BY BOB CARLSON, MHA
Senior Contributing Editor

In November 2005, author and Harvard alumna Laurie Levin had her lumpectomy and was diagnosed with stage 2 invasive ductal carcinoma. Then came the bad news.

The standard of care for this kind of breast cancer includes both radiation and chemotherapy, but Levin already had reached the recommended lifetime limits of certain chemotherapy agents during successful treatment for stage 3b histiocytic lymphoma in 1978.

Levin's oncologist recommended Onco*type* DX, a new genetic expression test that would quantify the likelihood of distant breast cancer recurrence. A low Onco*type* DX recurrence score also would correlate with a small likelihood of chemotherapy benefit.

"Once I knew I had a low recurrence score, everything lifted," Levin recalls. "Even before I finished the radiation treatments, I was already on the road to feeling okay about what had happened. No question in my mind, the Onco*type* DX test was absolutely critical for my recovery."

Levin is now one of more than 40,000 women who have had their tumor tissue analyzed by Genomic Health in Redwood City, Calif. Chief medical officer Steven Shak, MD, is gratified that Onco*type* DX allows breast cancer patients like Levin to make informed treatment decisions.

"That's why we do what we do," says Shak with a smile. He's smiling not only because of stories like Levin's, but because the American Society of Clinical Oncology (ASCO) and the National Comprehensive Cancer Network (NCCN) have included Onco*type* DX in their updated treatment guidelines.

In effect, the ASCO and NCCN recommendations establish Onco*type* DX as a new standard of care for a particular breast cancer patient population: Those whose disease is newly diagnosed, stage 1 or 2, node-negative, estrogen receptor (ER)-positive, and who will be treated with tamoxifen. This population accounts for about half of newly diagnosed breast cancer patients. Even though at least one similar test is commercially available, Onco*type* DX is the only one recommended by ASCO and NCCN under the new category of multiparameter gene expression analysis for breast cancer.

"The criteria being used by ASCO and other groups set an appropriately high bar for tests to be used during breast cancer treatment planning," Shak adds.

"DOING THE BEST THEY CAN"

Clinical studies published in such peer-reviewed journals as the *New England Journal of Medicine* and *The Lancet* have long been taken as unalloyed scientific fact. So would it be considered shocking if not every study published in these prestigious journals is first-rate?

"Alain Dupuy and I conducted a detailed review of about 45 papers in oncology research that were published in 2004, and related gene expression profiles to cancer outcomes for patients. We found that 50 percent of them had one or more of what we considered to be very se-

rious flaws," says Richard M. Simon, DSc, chief of the Biometric Research Branch in the Division of Cancer Treatment & Diagnosis at the National Cancer Institute.

Published in the *Journal of the National Cancer Institute*, the review focused on microarray use in cancer outcomes studies.[1] Because these microassays can measure the expression of as many as 35,000 genes, Simon cautions that their data require more sophisticated statistical techniques than studies that measure only one variable.

"Biologists and clinical investigators want to use technology like this because it's very powerful, but it's a real challenge for them because the data analysis is complicated," he says. "There are not enough statisticians available or necessarily knowledgeable on how to do this, and so they're sort of doing the best they can."

WHAT TO LOOK FOR IN A STUDY

But it's not just about the biostatistical challenges. In a 2006 article in *Clinical Advances in Hematology & Oncology*,[2] Simon included 17 "key issues" that may call study results into question. The following are among the most important:

Is it a developmental or validation study? Developmental and validation studies are akin to phase 2 and 3 clinical trials. For example, Onco*type* DX developmental studies identified the 21 genes (16 cancer related and 5 reference genes) whose expression appeared to have prognostic (cancer recurrence) and predictive (chemotherapy benefit) value, and gave a mathematical "weight" to each gene (the algorithm). Together, this generated the recurrence score (the classifier).

Validation studies corroborated that the 21-gene expression assay

[1] Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst.* 2007;99:147–157.

[2] Simon RM. Checklist for Evaluating Reports of Expression Profiling for Treatment Selection. *Clin Adv Hematol Oncol.* 2006;4:219–224.

Gary Wagner

"We basically took a lot of the good principles used in how to develop drugs and get the evidence to justify the use of a new drug and applied those same principles," says Steven Shak, MD, of Genomic Health.

Multiplexing will enable diagnoses based on a more informative assessment of biomarker panels, providing better disease prognosis and more effective patient management.

actually prognosticates and predicts what the developmental studies claimed. Ideally, validation studies use different patients and are conducted by researchers independent of those who performed the developmental studies.

"You really shouldn't use the same data to develop the classifier and then to evaluate it," says Simon. "Everybody loves to do developmental studies, creating new things with potential medical applications. A validation study is not quite so sexy — collecting a lot more data so that you can validate something that someone else did. The biggest caveat is that it's not ready for use with patients unless it's been validated in a separate study."

Are patients sufficiently homogeneous to be therapeutically relevant? Simon explains: "With tumor tissue specimens, what's often done is to use a department collection accumulated from patients with breast cancer over 10 or 20 years. However, these specimens usually are from a wide variety of patients who were treated in a wide variety of ways and whose cancer may have advanced to a wide variety of stages. Your new assay seems to be correlated with, say, disease-free survival. A lot of published oncology studies are of this type, and it's very hard to figure out what to do with that. It's important in doing these kinds of studies to have a therapeutically relevant question that can be answered and will help oncologists select treatments for patients. What oncologists really want to know is how to treat patients."

With Onco*type* DX, the therapeutically relevant question is "Does this node-negative, ER-positive patient have a sufficiently good prognosis on tamoxifen alone that the potential advantage from chemotherapy is minimal?" The recurrence score provides the answer.

Does the study provide information about assay reproducibility? "Genomic Health did the analytic validation of the Onco*type* DX assay," says Simon. "It's really technical reproducibility kind of stuff, showing that you can take a patient's tumor tissue and ship it across the country, and if you do the assay on different parts of the same tumor, you get the same answer, and if you do the assay twice in their lab, they get the same answer."

Focus on the importance of analytic validation was heightened after studies presented at the 2007 ASCO annual meeting reported inconsistent results from commercial HER2 assays of breast tumor tissue. HER2 tests identify responders to treatment with trastuzumab (Herceptin). ASCO addressed this issue in its guidelines by recommending that laboratories offering HER2 testing be annually accredited.

Shak contrasts the HER2 testing troubles with the Onco*type* DX assay.

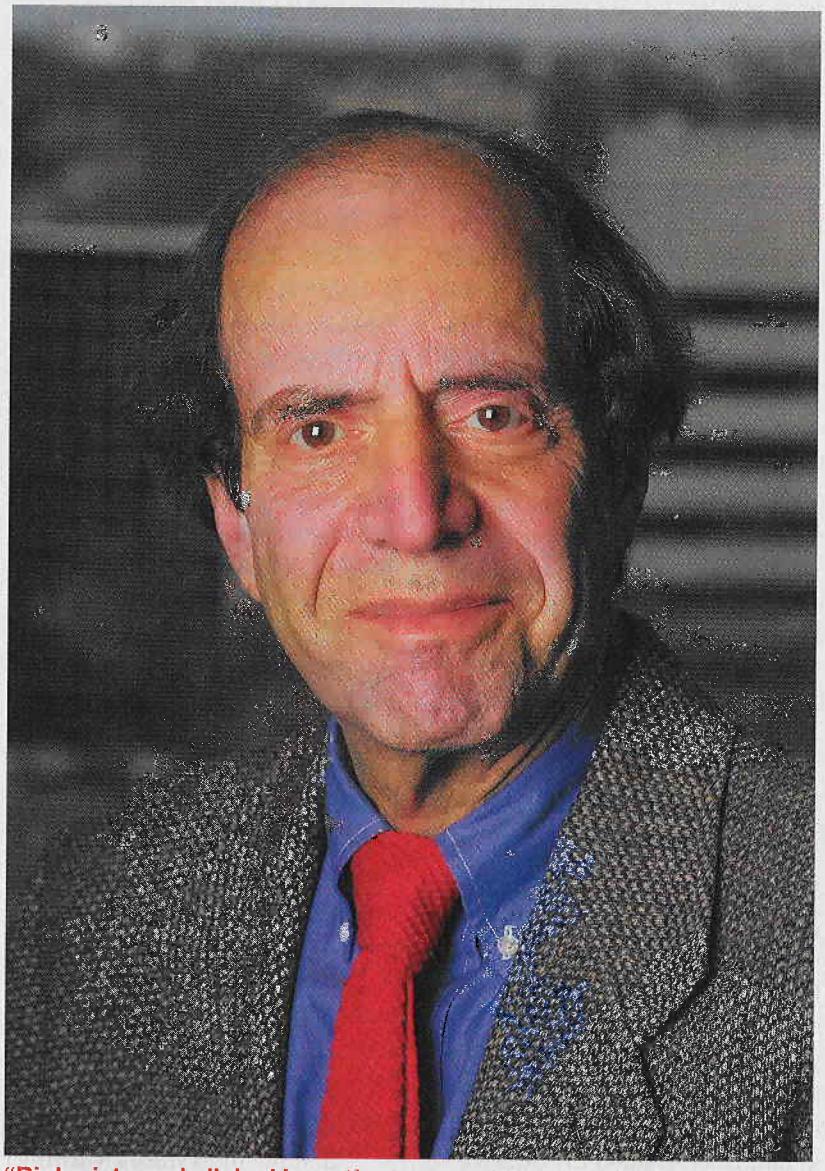
"We've made the investment, and continue to invest greatly, in standardizing our assay so that it can be run in a single lab where we get the same results again and again," he says. "That's so, so important."

EVIDENCE TO DEFINE QUALITY OF DATA

When developing a diagnostic for cancer treatment planning, "We basically took a lot of the good principles used in how to develop drugs and get the evidence to justify the use of a new drug and applied those same principles," says Shak. "We also listened to people like Richard Simon, who has outlined how to perform rigorous studies in the papers he's published."

Daniel F. Hayes, MD, clinical director of the Breast Cancer Oncology Program at the University of Michigan Comprehensive Cancer Center, Ann Arbor, was co-chair of the 2007 ASCO Tumor Markers Update Panel. Hayes, who recused himself from the deliberations regarding multigene assays such as Onco*type* DX because of his past association with Genomic Health as a paid consultant and active scientific collaborator, readily acknowledges Simon's influence. "A lot of what I have learned, I have learned from Richard," says Hayes.

A panel committee reviewed and analyzed data published since 1999 and evaluated 13 tumor markers, with some identified for multiple applications. The Multiparameter Gene Expression Analysis for Breast Cancer, which includes evaluations of the Breast Cancer Gene Expression Ratio; MammaPrint; and the Rotterdam Signature were new considerations for the 2007 guidelines.



Rob Crandall

"Biologists and clinical investigators want to use technology like this because it's very powerful, but it's a real challenge for them because the data analysis is complicated," says Richard M. Simon, DSc, of the National Cancer Institute.

"It's an enormous amount of work," Hayes says. "These aren't just opinion-based guidelines. We tried to make them as evidence-based as we could, to review all the available literature relative to the specific topic. Then we tried to use fair, balanced, and objective criteria for our recommendations."

One critical tool in this labor-intensive process was the Tumor Marker Utility Grading System, devised by ASCO panelists, including Hayes, when ASCO first decided to include tumor markers in its practice guidelines in 1995. The system established three Levels of Evidence to define the data quality on a given marker:

- Level of Evidence I, which is the most credible, designates data from prospective, randomized, controlled trials specifically designed to test the utility of the marker, or meta-analyses of well-designed studies.
- Level of Evidence II designates data from prospective therapeutic trials in which marker utility is a secondary study objective.
- Level of Evidence III designates data from large but retrospective studies.

It is this methodology, along with the biostatistical and clinical trial "dos and don'ts" that Simon writes about, that ultimately caused the ASCO panel to recommend Onco*type* DX, but not the Breast Cancer Gene Expression Ratio, MammaPrint, or the Rotterdam Signature.

Marketed by AviaraDx, in Carlsbad, Calif., the Breast Cancer Gene Expression Ratio quantifies the ratio of the HOXB6 and IL17BR genes in tumor tissue. Developmental studies were done with microarrays and frozen tissue, but validated by real-time quantitative polymerase chain reaction (RT-PCR) using the same tumor samples.

The ASCO Update Committee reports that the Breast Cancer Gene Expression Ratio "is significantly and independently associated with poorer disease-free survival in two studies of lymph node-negative, ER-positive, tamoxifen-treated patients with breast cancer." However, there are no published studies demonstrating that the assay is an improvement over conventional methods of classifying patients by recurrence outcomes, or of predicting chemotherapy benefit.

"Most of us just don't know how to use this [MammaPrint] assay," says Hayes. "It's been shown that this assay does indeed provide prognostic information, but it's not clear in what situation or how to apply that information to a specific patient. They have not actually addressed a specific clinical question."

A team at the Netherlands Cancer Institute in Amsterdam developed the 70-gene signature on which MammaPrint is based. Now marketed by Agendia, in Amsterdam, MammaPrint is a microarray-based prognostic assay for use in patients younger than age 53 with stage 1 or 2 primary, lymph node- negative breast cancer.

MammaPrint is the first assay cleared by the U.S. Food and Drug Administration as an in vitro diagnostic multivariate index assay device. Diagnostic device approvals are handled by the FDA's devices branch, the Center for Devices and Radiological Health, which is charged with verifying that a device is safe, analytically valid, and does what the sponsor claims. CDRH is not responsible, however, for demonstrating that use of the device improves clinical outcomes.

"MammaPrint does not have FDA approval," Hayes emphasizes. "Agendia has FDA clearance to sell MammaPrint in the United States. When we deliberate in ASCO, we aren't influenced by whether the FDA has cleared a marker, because that doesn't indicate a change in patient outcomes at all. The devices branch has a very different set of criteria than the [Center for Drug Evaluation and Research], which approves a drug as safe an effective."

In other words, ASCO evaluated MammaPrint strictly on its data quality and the design of the clinical studies that generated that data.

"In some of those studies, some patients were treated with systemic therapy and some were not," Hayes says. "In the so-called 'validation study,' they actually used a percentage of the samples from their development set."

Hayes also is cautious about microarray-based gene expression signatures, but for a different reason than Simon.

"Onco*type* DX was developed not with microarrays but with RT-PCR," Hayes explains. "One reason that's important is that there is concern that those arrays have not been technically validated in regard to reproducibility. I have not seen a paper showing that you can take the same sample and get the same result three times, although FDA clearance suggests that these data must have been generated."

Tumor marker research and development efforts

- The Breast Cancer Intergroup is conducting a study to evaluate whether women with mid-range Onco*type* DX Recurrence Scores benefit from chemotherapy.
- According to Wang, at Veridex, a multi-center study validating the GeneSearch Breast Prognostic Assay in tamoxifen-treated ER-positive patients has been completed but not yet published.
- Multiple study results presented at December's San Antonio Breast Cancer Symposium suggest that Onco*type* DX may be clinically useful for certain node-positive patients treated with chemohormonal therapy followed by tamoxifen, and that estrogen receptor (ER) expression may be prognostic in ER-positive and ER-negative patients.
- Genomic Health is working on genetic expression assays for colon, prostate, renal, and lung cancers, along with melanoma.

GENESEARCH BREAST PROGNOSTIC ASSAY

The 76-gene microarray-based Rotterdam Signature was developed at the Daniel den Hoed Cancer Center, Erasmus Medical Center, in Rotterdam. Not yet commercially available, the Rotterdam Signature is being developed by Veridex, a Johnson & Johnson company.

"When we defined this 76-gene signature, we used retrospective patient cohorts who had received surgery alone, without adjuvant chemotherapy or hormonal therapy, because we wanted to define a pure prognostic test that can accurately and reliably determine the patients with low and high risk of disease recurrence," says Yixin Wang, PhD, executive director of research and development at Veridex, in Warren, N.J., and an early collaborator with the Rotterdam team. Veridex is developing the Rotterdam Signature under the name of GeneSearch Breast Prognostic Assay.

In a 2005 article in *The Lancet*, Wang and colleagues reported that the 76-gene expression profile identified patients who went on to develop distant metastases within 5 years with 93 percent sensitivity and 48 percent specificity. Multicenter validation studies of the 76-gene signature also appeared in the *Journal of Clinical Oncology* and *Clinical Cancer Research*.

In a clear reference to Onco*type* DX, a Veridex news release announcing the 2005 *Lancet* paper noted "currently available tools are generally restricted to patients with a specific ER status, or to patients already taking tamoxifen." That may not be the only po-

"It's a field that's full of both hype and nonsense, but at the same time, there's a lot of wonderful substance there," says Richard M. Simon, DSc, at the NCI.

tential advantage of the GeneSearch Breast Prognostic Assay, which was developed and validated in lymph node-negative breast cancer patients regardless of age, tumor size, grade, menopausal, and ER status. Wang states that the assay's 90 to 95 percent sensitivity rate compares favorably to other tests, but, more importantly, distinguishes breast tumors with a very low risk of recurrence, 5 percent or less, versus those with a much higher risk, between 30 and 40 percent. It may offer a reliable way to identify patients who can be cured by surgery alone without aggressive and toxic chemotherapy, he says.

Wang attributes at least part of the difference in test performance to the "whole genome" development approach versus the "targeted gene" development approach that some other tests use. "In the discovery data published in *The Lancet*, we pointed out that among the 76 genes in the prognostic signature, there are 18 novel genes about which we don't have prior knowledge of their biological function," says Wang. "In other words, this data highlights the power of the whole genome approach to identify not only the genes that have been implicated in breast cancer, but also new genes whose relationship to the underlying disease is not yet understood. That's probably part of the reason why this kind of gene signature performed better than some of the other tests."

WONDERFUL SUBSTANCE

Most payers have dedicated nurses and physicians who track and evaluate new technologies for which coverage policies may have to be developed. Their evaluation process is similar to how the ASCO Update Committee reached its recommendations, such as determining whether the clinical trials are well designed and the biostatistical analyses correctly performed.

"This is a huge issue because there are very few truly double-blinded studies that allow us to make firm decisions on a new technology," says Louis I. Hochheiser, MD, medical director of clinical policy development at Humana. "Molecular diagnostics companies are trying to bring their products to market as quickly as possible, and the level of evidence that you really would like to have is oftentimes not there."

Which is why Hochheiser and colleague Bryan Loy, MD, MBA, medical officer in Humana's Kentucky market office, welcome ASCO's evaluations and recommendations. In the interim, Hochheiser says Humana is hiring more people to focus solely on the fast-growing molecular diagnostics category.

The upside of these new gene expression assays for payers could be lower total costs. Data from 2005 estimated that compared with treatment decisions based on NCCN criteria, clinical decisions based on Onco*type* DX results could boost average quality-adjusted survival by 8.6 years and reduce overall costs by $202,828 for 100 theoretical U.S. patients. Still, many payers are concerned about adopting this new technology indiscriminately.

"This field has exploded in the last 6 to 9 months with a high level of awareness and a lot of promise about where things can go," says Hochheiser. "But it's being pushed so fast that the medical evidence may not yet support its usefulness in the clinical setting. We need to ensure that the medical evidence is there, enabling us to appropriately take care of people's medical problems."

Humana, though, has made an early decision, based on medical evidence, to adopt Onco*type* testing. Patients should have the access to the latest technology, Hochheiser believes, as long as they are protected from harm.

Simon at the NCI also encourages buyers and payers alike to beware of what they hear and read about molecular diagnostics. He agrees that it will take time before we understand tumor biology better and why tests like Onco*type* DX work, but he doesn't think this is a time to "step back."

"It's a field that's full of both hype and nonsense, but at the same time, there's a lot of wonderful substance there," says Simon. "To imply that this technology isn't mature and not to be trusted I think is erroneous. Trying to understand the biology of a disease is a lifetime's worth of work. Let's not make the mistake of saying we're not going to accept the use of such a test unless we understand why it works biologically."

Laurie Levin and tens of thousands of other women, one suspects, would agree.

Bob Carlson, MHA, writes exclusively about healthcare. He lives near Zionsville, Ind.

Biotechnology Healthcare

THE IMPACT OF BIOLOGICS ON HEALTH, BUSINESS, AND BENEFITS

MAY/JUNE 2008

www.biotechnologyhealthcare.com

FEATURE ARTICLES

FOCUS: ONCOLOGY

DEPARTMENTS

BIOTECHNOLOGY HEALTHCARE (ISSN 1554-169x) is published bimonthly by BioCommunications LLC, 780 Township Line Road, Yardley, PA 19067. This is Volume 5, Number 1. Standard postage paid at Morrisville, Pa., and at additional mailing offices. Price: $18 per copy, $95 per year in the United States; $120 per year elsewhere. Phone: (267) 685-2788; fax (267) 685-2966; paid subscription inquiries (267) 685-2789. Letters to the editor are accepted for publication consideration at the above address.

INDEXED BY CINAHL AND INTERNATIONAL PHARMACEUTICAL ABSTRACTS

COVER: JOHN STILL /GETTY IMAGES