MOLECULAR ORIGINS OF CANCER

Gene-Expression Signatures in Breast Cancer

Christos Sotiriou, M.D., D.Phil., and Lajos Pusztai, M.D., D.Phil.

G
ene-expression profiling with the use of DNA microarrays allows measurement of thousands of messenger RNA (mRNA) transcripts in a single experiment. Results of such studies have confirmed that breast cancer is not a single disease with variable morphologic features and biomarkers but, rather, a group of molecularly distinct neoplastic disorders. Profiling results also support the hypothesis that estrogen-receptor (ER)–negative and ER-positive breast cancers originate from distinct cell types and point to biologic processes that govern metastatic progression. Moreover, such profiling has uncovered molecular signatures that could influence clinical care. In this review, we summarize the results of gene-expression studies that hold the most promise to accelerate the transition between empirical and molecular medicine.

MOLECULAR CLASSIFICATION OF BREAST CANCER

Four main molecular classes of breast cancer have been distinguished by gene-expression profiling.

1-5 The “intrinsic” classification by Perou et al.1 proposes that these four classes be called basal-like breast cancers, which mostly correspond to ER-negative, progesterone-receptor (PR)–negative, and HER2-negative tumors (hence, “triple-negative” tumors); luminal-A cancers, which are mostly ER-positive and histologically low-grade; luminal-B cancers, which are also mostly ER-positive but may express low levels of hormone receptors and are often high-grade; and HER2-positive cancers, which show amplification and high expression of the ERBB2 gene and several other genes of the ERBB2 amplicon. These subgroups correspond reasonably well to clinical characterization on the basis of ER and HER2 status, as well as proliferation markers or histologic grade.

Microarray studies have shown that luminal types of tumors express high amounts of luminal cytokeratins and genetic markers of luminal epithelial cells of normal breast tissue.6 In contrast, basal-like breast cancers do not express ER, PR, and ER-related genes and do not overexpress several genes that typify myoepithelial cells of normal breast tissue: luminal cytokeratins, smooth-muscle–specific markers, and certain integrins. In some basal-like cancers, there is high expression of “basal” cytokeratins such as CK5 and a variety of growth factor receptors, including high levels of epidermal growth factor receptor, c-kit (a tyrosine kinase in breast epithelium), and growth factors such as hepatocyte growth factor and insulin growth factor.3-4 Immunohistochemical methods for defining basal-like cancers7 have not gained wide acceptance, partly because correspondence with molecular classification is less than perfect and also because logistic complexities limit the feasibility of combining five or more immunohistochemical markers in routine clinical practice.

Another feature that differentiates sporadic basal-like tumors from luminal-like tumors is dysfunction of the BRCA1 pathway caused by BRCA1 gene promoter methyl-
BRCA1 expression is important in DNA repair, activation of cell-cycle checkpoints, maintenance of chromosomal stability, and perhaps differentiation of ER-negative stem or progenitor cells into ER-positive luminal cells. These findings are in line with suggestions of a link between the basal-like phenotype and germ-line mutation of BRCA1. Indeed, almost all breast cancers that are associated with a BRCA1 mutation, whether sporadic or hereditary, have a basal-like triple-negative phenotype. Tumors associated with the BRCA2 mutation have the distribution of phenotypes encountered in the general population.

Tumor grade can discriminate luminal A from luminal B tumors. This distinction can be further refined by the application of a genomic grade, a gene-expression signature of tumor differentiation. Luminal B tumors typically have a high genomic grade, similar to basal-like and HER2-positive tumors, whereas luminal A tumors have a genomic grade similar to that of normal breast tissue.

Microarray-based comparative genomic hybridization has revealed differences in copy numbers of particular genes in different subtypes of breast cancer. The increased copy-number variation in basal-like tumors indicates more genetic complexity than in the other subtypes, suggesting a greater degree of genetic instability in these tumors. Basal-like cancers are relatively enriched for low-level copy-number gains involving several chromosomal regions, whereas high-level amplification at any locus is infrequent. In contrast, high-level amplifications are seen more frequently in HER2-positive and luminal B tumors. Similar aberrant genomic patterns occur in familial breast cancers that are not associated with BRCA1 or BRCA2. Both hereditary BRCA1-associated tumors and sporadic basal-like tumors do not have markers of X-chromosome inactivation (Xi); duplication of the active X chromosome and loss of Xi suggest that X-chromosome abnormalities contribute to the pathogenesis of basal-like cancers.

These distinct transcriptional and genomic aberrations that differentiate the four subtypes of breast cancer indicate that these variants may arise from different transformed stem or progenitor cells, each with distinct biologic properties. Moreover, these subgroups track with prognosis and responses to therapy. The low-grade luminal A tumors are indolent and sensitive to antiestrogens. Luminal B tumors and tumors that are HER2-positive and ER-positive have incomplete sensitivity to endocrine therapy, and HER2-positive tumors, which have an aggressive natural history, are sensitive to trastuzumab, an anti-HER2 antibody. Basal-like tumors also have a more aggressive natural history, though they can be especially sensitive to chemotherapy.

The additional clinical value of molecular classification is limited by its close correspondence with the status of ER, PR, and HER2, along with tumor grade (Fig. 1). However, molecular classification is changing the design of clinical trials. Moreover, the molecular differences that underlie the phenotypes of breast cancer could reveal new therapeutic targets. Examples are the identification of a functional androgen-receptor pathway in a subgroup of ER-negative and PR-negative breast tumors and defects in DNA-repair pathways in BRCA1 and BRCA2 carriers and probably in many basal-like cancers.

**Gene-expression Signatures and Clinical Outcome**

Gene-expression profiling has been used to develop genomic tests that may provide better predictions of clinical outcome than the traditional clinical and pathological standards (Table 1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). Three different strategies have been explored for this purpose (Fig. 2).

Using the supervised top-down approach, investigators from the Netherlands Cancer Institute developed a gene signature (MammaPrint, Agendia) from a selected retrospective series of 78 patients with node-negative breast cancer who had received no systemic adjuvant therapy. The assay, which measures the expression of 70 genes and calculates a prognostic score that categorizes patients into “good” or “poor” risk groups, was recently cleared by the Food and Drug Administration (FDA) to aid in formulating a prognosis for patients with breast cancer who are under 61 years of age and who have node-negative, stage I or II disease with a tumor size of 5 cm or less. However, the assay has not been tested in a prospective study. The Dutch researchers also reported a validation study from a retrospectively collected consecutive series of breast tumors, including both node-negative and node-positive cancers. In this study, however, 130 patients had received systemic adjuvant chemotherapy or hor-
monal therapy, and 61 had also been included in the original study. Therefore, the results could have been biased. A second, more appropriate validation study that included 307 patients who had received no systemic therapy confirmed the Dutch findings.

A comparison of this gene signature with the Adjuvant! Online program (www.adjuvantonline.com), which assigns risk according to conventional criteria of tumor size, nodal status, grade, and ER status, showed that 87 of 302 patients had discordant results (29%). Of these 87 patients, 59 (68%) had tumors that were rated as clinically high-risk according to the conventional criteria but low-risk according to their gene signature, and 28 (32%) had tumors that were rated as clinically low-risk but high-risk according to their gene signature. In these discordant cases, the genomic test appeared to predict the outcome more accurately. Patients with clinically low-risk tumors that were rated as high-risk on genomic assay had a 10-year overall survival rate of 69%, whereas patients with clinically high-risk tumors that were low-risk according to their gene signature had a 10-year overall survival rate of 89%. Whether these data would be relevant in patients treated with adjuvant therapy remains unclear. It is interesting that in a prospectively conducted multicenter study...
that included 427 patients in various Dutch hospitals, the use of MammaPrint in combination with clinical guidelines led to altered adjuvant treatment recommendations in 26% of patients.46

The genomic-grade signature exemplifies the “bottom-up” discovery strategy.39 It seeks to define molecular features of tumor differentiation and tumor grade, both of which influence tumor progression and metastatic spread.47-49 A 97-gene signature consistently discriminated between low-grade and high-grade tumors. This signature, which is driven by proliferation and cell-cycle genes, separates the intermediate-grade tumors that are problematic for making decisions about treatment into two subgroups of low genomic grade and high genomic grade, with outcomes similar to those of low and high histologic grade, respectively. These results were observed across multiple independent data sets that were generated on different microarray platforms. In more than 650 patients with ER-positive breast cancer who were untreated or who had received only tamoxifen, genomic grade was associated with outcome more than clinical variables.16 This result highlights the importance of tumor-differentiation and tumor-proliferation genes in the ER-positive subgroup, as reported previously.16,37,50-52

MammaPrint, the genomic-grade signature, and a 76-gene outcome signature (developed by researchers from Rotterdam, the Netherlands, in collaboration with Veridex) appear to quantify mainly tumor grade and proliferation. When the three assays were analyzed in the same population of patients who had received no systemic adjuvant therapy, they had similar performance, suggesting that genes controlling tumor differentiation and proliferation account for a large proportion of these classifiers.53

Another molecular assay, Oncotype DX (Genomic Health), exemplifies the candidate-gene approach to estimating outcome (Table 1).43 It measures the expression of ER and HER2, as well as that of ER-regulated transcripts and several proliferation-related genes, with the use of the quantitative reverse-transcriptase–polymerase-chain-reaction (RT-PCR) assay. Most of these genes are associated with outcome, and several can be assessed with the use of conventional methods. The Oncotype DX system combines these measurements into a quantitative “recurrence score,” which can be used as a continuous variable to estimate the probability of recurrence at 10 years or to group patients into low-risk, intermediate-risk, and high-risk categories.

The association between the recurrence score and distant relapse was examined retrospectively in 668 patients with ER-positive, node-negative cancers treated with tamoxifen who were enrolled in the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14 clinical trial. The 10-year distant recurrence rates were 7%, 14%, and 30% for the low-risk, intermediate-risk, and high-risk categories, respectively. Similar results were found in a community-based population of patients.55 The Oncotype DX assay appears to identify tumors that are likely to respond to adjuvant chemotherapy in addition to tamoxifen...
therapy. The association of the recurrence score with benefit from adjuvant cyclophosphamide, methotrexate, and fluorouracil chemotherapy in ER-positive, node-negative, tamoxifen-treated patients was examined in 651 patients who were enrolled in the NSABP B-20 randomized clinical trial. Higher recurrence scores were associated with greater benefit from adjuvant chemotherapy, and more critically, lower recurrence scores were associated with a lack of even marginal benefit from chemotherapy. Similar results were found in a subgroup analysis of the Southwest Oncology Group (SWOG) Intergroup 0100 trial, a randomized study of tamoxifen with or without anthracycline-based chemotherapy for postmenopausal women with node-positive breast cancer.

In limited feasibility studies, published only in abstract form, it was reported that molecular information may change treatment recommendations for approximately 30% of patients, usually prompting less use of chemotherapy. The Oncotype DX assay has been endorsed as a tumor marker by the American Society of Clinical Oncology (ASCO) and as an aid to decision making regarding adjuvant chemotherapy in patients with ER-positive, node-negative breast cancer by the breast cancer panel of the National Comprehensive Cancer Network (NCCN) (www.nccn.org).

Other gene signatures that may predict the risk of recurrence in ER-positive patients treated with tamoxifen have been developed through a “bottom-up” discovery strategy (Table 2 in the Supplementary Appendix). Among these signatures, a high mRNA-expression ratio of HOXB13 to IL17R (H/I) was associated with a high risk of recurrence in patients treated with tamoxifen. The H/I ratio was also confirmed in independent retrospective series with the use of standard formalin-fixed, paraffin-embedded tissue samples from both untreated and tamoxifen-treated patients. Recently, the accuracy of the H/I assay was improved by including the molecular-grade index, which mainly measures cell proliferation. The fact that different gene signatures have very few genes in common may be surprising at first, but it is a common feature of complex gene-expression profiling.

Table 2. Commercially Available Genomic Assays for the Prediction of Clinical Outcome in Patients with Breast Cancer.6

<table>
<thead>
<tr>
<th>Variable</th>
<th>MammaPrint</th>
<th>Oncotype DX</th>
<th>Theros</th>
<th>MapQuant Dx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider</td>
<td>Agendia</td>
<td>Genomic Health</td>
<td>Biotheranostics</td>
<td>Ipsogen</td>
</tr>
<tr>
<td>Type of assay</td>
<td>70-Gene assay</td>
<td>21-Gene recurrence score</td>
<td>2-Gene ratio of HOXB13 to IL17R (H/I) and molecular-grade index</td>
<td>Genomic grade</td>
</tr>
<tr>
<td>Type of tissue sample</td>
<td>Fresh or frozen</td>
<td>Formalin-fixed, paraffin-embedded</td>
<td>Formalin-fixed, paraffin-embedded</td>
<td>Fresh or frozen</td>
</tr>
<tr>
<td>Technique</td>
<td>DNA microarrays</td>
<td>Q-RT-PCR</td>
<td>Q-RT-PCR</td>
<td>DNA microarrays</td>
</tr>
<tr>
<td>Centrally certified laboratory†</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Indication</td>
<td>To aid in prognostic prediction in patients &lt;61 yr of age with stage I or II, node-negative disease with a tumor size of ≤5 cm</td>
<td>To predict the risk of recurrence in patients with ER-positive, node-negative disease treated with tamoxifen; to identify patients with a low risk of recurrence who may not need adjuvant chemotherapy</td>
<td>To stratify ER-positive patients into groups with a predicted low risk or high risk of recurrence and a predicted good or poor response to endocrine therapy</td>
<td>To reclassify grade 2 tumors into low-risk grade 1 or high-risk grade 3 tumors, specifically for invasive, primary, ER-positive grade 2 tumors</td>
</tr>
<tr>
<td>Level of evidence (I–V)‡</td>
<td>III</td>
<td>II</td>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>FDA clearance</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Availability</td>
<td>Europe and United States</td>
<td>Europe and United States</td>
<td>United States</td>
<td>Europe</td>
</tr>
</tbody>
</table>

ER denotes estrogen receptor, FDA Food and Drug Administration, and Q-RT-PCR quantitative reverse-transcriptase–polymerase chain reaction.

† Laboratories were certified according to the criteria of the Clinical Laboratory Improvement Amendments or by the International Organization for Standardization.

‡ Levels of evidence are measured on a scale ranging from I (strongest) to V (weakest).
expression data that contain large numbers of highly correlated variables (i.e., gene-expression measurements). Several different combinations of the correlated variables can be selected to build similarly accurate prediction models. Indeed, in a study in which five different signatures were tested on the same data, four of the five had similar performance. This study also showed that the different signatures identified a largely overlapping population of patients as high-risk. Among the ER-positive, luminal B tumors, the rates of high-risk designation were 93% with the use of Oncotype DX and 84% with the use of MammaPrint. This concordant risk assignment occurred even though only a single gene (SCUBE2) is common to MammaPrint and Oncotype DX. An important limitation of these classifiers is that they assign the high-risk category to almost all ER-negative patients.

A meta-analysis of publicly available gene-expression and clinical data from almost 3000 breast tumors supported the relationship between the risk of recurrence and molecular subtype, including several different signatures, as well as routine clinical and pathological variables. An encouraging finding is that all evaluated signatures showed similar performance despite the limited overlap of genes. Several common features also emerged from this analysis: basal-like or triple-negative tumors and HER2-positive tumors had high expression of tumor-differentiation genes, including several cell-cycle and proliferation genes. In contrast, the ER-positive combined luminal A and luminal B subtypes were more heterogeneous. The low-risk luminal A tumors were associated with a low expression of proliferation-related genes. The expression of several cell-cycle and proliferation-related genes drove the performance of several signatures. All the signatures were most useful in ER-positive tumors as a consequence of identifying the low-proliferation luminal A tumors at low risk of recurrence, whereas they were less informative for the basal-like and HER2-positive tumors, since most of these tumors were classified as high-risk by all signatures. Testing with more than one signature did not improve the performance, and lymph-node status and tumor size, which essentially capture the clinical tumor stage, had an independent prognostic value. These results suggest that both genomic and clinical variables should be included in a common algorithm to yield the most accurate prediction model (Fig. 3).

Other gene signatures that assess the role of tumor microenvironment, chromosomal instability, stem-cell biology, and the process of metastatic spread and colonization have been reported. Although they yield new insights into the biology of stem cells and the metastatic process, their clinical use beyond that of other classifiers is unknown.

**Gene-expression signatures and response to chemotherapy**

The development of tests to predict responses to chemotherapy poses several practical challenges. There are theoretical limits to the accuracy of any response predictor that measures the characteristics of only the cancer. Host characteristics, including the rate of drug metabolism, and the association between benefit from tamoxifen and the genetic variants of CYP2D6, a cytochrome involved in the metabolism of tamoxifen, can also affect the response to therapy. Moreover, there is considerable uncertainty as to what level of predictive accuracy would be clinically useful, since different levels of accuracy may be required for different clinical situations, depending on the availability of alternative treatments, the frequency and severity of adverse effects, and the risks of disease progression in the absence of therapy. Given these complexities, many of the studies of genomic markers have focused on preoperative (neoadjuvant) treatment in breast cancer. Pathologic complete response to chemotherapy indicates that the cancer is extremely sensitive to chemotherapy. Most of the clinical trials examining the correlation between pathologic complete response and long-term cancer-free survival have reported a strong association between these two outcomes.

Several small studies have shown that the gene-expression profiles of cancers that are highly sensitive to chemotherapy differ from those of less responsive tumors. The largest study to date included prospectively collected needle-biopsy samples from 133 patients with stage I, II, or III breast cancer who received preoperative weekly paclitaxel and a combination of fluorouracil, doxorubicin, and cyclophosphamide. Data from the first 82 patients were used to develop a multigene signature predictive of pathologic complete response, and data from the remaining 51 patients were used to test the accuracy of the predictor. This 30-gene predictor showed higher sensitivity than a clinical predictor that included age, nuclear grade, and...
ER status (92% vs. 61%). It also correctly identified 92% of the patients who achieved a pathologic complete response. The positive predictive value of the pharmacogenomic signature was a modest 52%, but its negative predictive value was 96%. Similar results were reported in several other small pilot studies (Table 3 in the Supplementary Appendix). None of these predictors of chemotherapy response are commercially available, and much larger studies are needed to validate these observations, assess the specificity of the treatment regimen, and determine the true performance characteristics of these tests.

An alternative approach is the use of experimental cancer models to define gene signatures that correlate with the response to particular drugs and to test the performance of these signatures in humans. A few groups have reported gene-expression signatures associated with response or resistance to chemotherapy in vitro.\(^78^{–}80\) Initial efforts to validate genomic predictors that are derived from cell lines with the use of data from humans have been reported,\(^81\) but the results remain unconfirmed.

**Figure 3. Molecular Classification, Gene-Expression Signatures, and Clinical Outcome.**

Genes that are associated with tumor differentiation and cell cycle drive the prognostic power of the intrinsic molecular classification and several gene-expression signatures.

Several signatures are under clinical development, and some that are commercially available have been cleared by the FDA for clinical use (i.e., MammaPrint) or endorsed by ASCO and NCCN guidelines (i.e., Oncotype DX) to assist clinicians in making decisions about treatment (Table 1).\(^82,83\) However, appropriate treatment recommendations can often be made without using these tests. The genetic tests add modest prognostic information for patients with HER2-positive and triple-negative tumors, but when measures of clinical risk are equivocal (e.g., intermediate expression of ER and intermediate histologic grade), these assays could guide clinical decisions.

Although many studies document the repro-
ducibility and associative properties of these assays, it is noteworthy that no prospective, randomized comparisons between genomic testing and clinical factors in making clinical decisions have been completed. Two such studies are under way: the Microarray in Node Negative and 0 to 3 Positive Lymph Node Disease May Avoid Chemotherapy Trial (MINDACT) in Europe\(^84\) (testing MammaPrint) and the Trial Assigning Individualized Options for Treatment (TAILORx) study in the United States\(^85\) (testing Oncotype DX) (Table 2). The results of these trials could provide valuable information about the use of gene-expression signatures in daily breast-cancer management and should address important logistic, technical,\(^87,88\) and analytical\(^89\) issues, such as those related to the handling, shipping, reproducibility, quality control, and standardization of these new molecular tools.

### Table 2. Key Features of the TAILORx and MINDACT Validation Trials.\(^*\)

<table>
<thead>
<tr>
<th>Variable</th>
<th>TAILORx</th>
<th>MINDACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligibility</td>
<td>Node-negative, ER-positive or PR-positive, HER2-normal, stage I or II tumor</td>
<td>Node-negative and 0 to 3 node-positive, any hormone receptor status, stage I, II, or III tumor</td>
</tr>
<tr>
<td>Molecular assay</td>
<td>Oncotype DX (Q-RT-PCR)</td>
<td>MammaPrint (DNA microarray)</td>
</tr>
<tr>
<td>Tissue requirement</td>
<td>Formalin-fixed, paraffin-embedded</td>
<td>Fresh or frozen</td>
</tr>
<tr>
<td>Number of participants</td>
<td>10,500</td>
<td>6000</td>
</tr>
<tr>
<td>Number to be randomized</td>
<td>4,390</td>
<td>1920</td>
</tr>
<tr>
<td>Randomized group</td>
<td>Patients with a recurrence score of 11 to 25 (44%)(†)</td>
<td>Discordant risk between AdjuvantOnline and MammaPrint (32%)(†)</td>
</tr>
<tr>
<td>Randomization</td>
<td>Endocrine therapy alone or endocrine therapy plus chemotherapy</td>
<td>Treatment recommendation on the basis of clinical or genomic risk</td>
</tr>
<tr>
<td>Treatment of nonrandomized groups</td>
<td>For recurrence score of &lt;11 (25%), endocrine therapy; for recurrence score of &gt;25 (27%), chemotherapy plus endocrine therapy(†)</td>
<td>For low risk on both predictions (13%), endocrine therapy alone; for high risk on both predictions (55%), chemotherapy plus hormonal therapy(†)</td>
</tr>
<tr>
<td>Primary research question</td>
<td>To determine whether adjuvant endocrine therapy alone is not inferior to chemotherapy and endocrine therapy in patients with an intermediate recurrence score (11 to 25)</td>
<td>To determine whether chemotherapy can be safely avoided in patients who are predicted to be at low risk by MammaPrint but at high risk by AdjuvantOnline</td>
</tr>
<tr>
<td>Primary end point</td>
<td>Disease-free survival</td>
<td>Distant metastasis-free survival</td>
</tr>
<tr>
<td>Secondary objectives</td>
<td>To create a tissue bank (formalin-fixed, paraffin-embedded samples and blood)</td>
<td>To create a specimen and gene-expression data bank (frozen or formalin-fixed, paraffin-embedded samples and blood)</td>
</tr>
<tr>
<td>Risk groups</td>
<td>Risk of distant recurrence in patients with a recurrence score of 11 to 25 treated with hormonal therapy, about 12.5%</td>
<td>Risk of death at 10 yr for low-risk patients: untreated, about 12%; treated with endocrine therapy, &lt;10%</td>
</tr>
<tr>
<td>Treatment regimens</td>
<td>Choice of treating physician for both endocrine and chemotherapy</td>
<td>For endocrine therapy for ER-positive tumors: randomization between tamoxifen for 2 yr, followed by letrozole for 5 yr or letrozole for 7 yr; for chemotherapy: randomization between anthracycline-based regimen or docetaxel–capecitabine</td>
</tr>
</tbody>
</table>

\(\ast\) The Microarray in Node Negative and 0 to 3 Positive Lymph Node Disease May Avoid Chemotherapy Trial (MINDACT) (ClinicalTrials.gov number, NCT00433589) was coordinated by the European Organization for Research and Treatment of Cancer, and the Trial Assigning Individualized Options for Treatment (TAILORx) (NCT00310180) was coordinated by the Eastern Cooperative Oncology Group. ER denotes estrogen receptor, PR progesterone receptor, and Q-RT-PCR quantitative reverse-transcriptase–polymerase chain reaction.

\(\dagger\) Percentages denote the proportions of patients in the study.

**Conclusions**

Results from studies of gene-expression profiling have altered our view of breast cancer and provided us with a new tool for molecular diagnosis. Technical advances are rapid in this field, and the
microarray platforms that were used to develop these signatures interrogate the “mRNA world.” The next generation of DNA microarrays (e.g., tiling arrays, microRNA arrays, and direct sequencing of complementary DNA) will enable investigators to study the clinical and diagnostic potential of new RNA species, including microRNAs and RNA transcribed from noncoding DNA, pseudogenes, and antisense DNA strands.

An exciting prospect of microarray-based tests is that multiple, distinct predictions—including prognosis, ER, and HER2 status, and sensitivity to various treatment approaches—could be generated from a single assay. This type of test would use information from different sets of genes from the same tissue for different predictions. This outcome is technically feasible and could substantially improve the cost-effectiveness of a multigene assay. To provide treatment recommendations that are truly molecularly tailored to individual patients in the future, it will be important to measure the risk of relapse and the probability of benefit from endocrine therapy and chemotherapy separately and to consider the preferences of patients in light of these results. Another promising direction of research is to examine the hypothesis that different markers and biologic pathways may be involved in determining prognosis, response, and resistance to therapy in different molecular subgroups of breast cancers. As ever-larger clinical data sets become available for gene-expression analysis, it is conceivable that predictors of molecular-class-specific prognosis and treatment response will be developed in the future.

Supported by grants from the Breast Cancer Research Foundation (to Drs. Sotiriou and Puotinen) and the Belgian National Foundation for Research (to Dr. Sotiriou).

No potential conflict of interest relevant to this article was reported.

We thank Drs. Martine Piccart-Gebhart and Gabriel N. Hortobagyi for their comments.

REFERENCES


