Mol Diag Ther 2009; 13 (2): 73-90
1177-1062/09/0002-0073/$49.95/0

© 2009 Adis Data Information BV. All rights reserved.

Genomic Predictors of Outcome and Treatment Response in Breast Cancer

Lara Dunn and *Angela DeMichele*

Abramson Cancer Center, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Contents

Abstract 73
1. Prognostic Gene Expression Signatures in Breast Cancer 75
   1.1 Molecular Subtypes Using the Intrinsic Gene Set ('Intrinsic Subtypes'). 75
   1.2 Oncotype DX® 21-Gene Profile Recurrence Score. 75
   1.3 MammaPrint® 70-Gene Profile 81
   1.4 The *HOXB13/IL17BR* Ratio (H/I Test) 81
   1.5 Rotterdam Signature. 82
   1.6 Wound Response Indicator 82
   1.7 Invasive Gene Signature 82
   1.8 Comparison of Prognostic Signatures. 82
2. Molecular Profiles Predicting Responsiveness to Chemotherapy 83
   2.1 Oncotype DX® 83
   2.2 MammaPrint® 85
   2.3 The H/I Test 85
3. Profiles Predicting Response to Specific Chemotherapeutic Agents. 85
4. Conclusions 88

Abstract

Despite advances in breast cancer treatment and outcome over the last two decades, women continue to relapse and die of advanced disease. Historically, estrogen and progesterone receptor expression, HER2 overexpression and clinico-pathologic parameters have guided therapeutic decision making. However, there are limits to the risk estimation provided by these parameters, leading to potential overtreatment of low-risk disease and undertreatment of poor-risk disease. Genomic technologies now provide the opportunity to refine our therapeutic approach by individualizing treatment to patients' individual tumor profiles. Gene profiles or signatures are groupings of genes that are differentially expressed between tumors, reflecting differences in biologic behavior. Prognostic gene signatures stratify breast cancer patients by tumor natural history, regardless of the treatment employed. Currently, there are three commercially available prognostic gene signatures: Oncotype DX® (Genomic Health, Inc.), MammaPrint® (Agendia BV), and the *HOXB13/IL17BR* (H/I) ratio; (Theros H/I^{SM}; bioTheranostics). Others under development include the Intrinsic Gene Set, the Rotterdam Signature, the Wound Response Indicator, and the Invasive Gene Signature. Predicative signatures classify patients based on responsiveness to specific therapies. Of the prognostic signatures, Oncotype DX® has been shown to have predictive value for the incremental benefit of chemotherapy when added to a hormonal therapy regimen. Additional genetic profiles under development predict response to specific hormonal therapies, anthracyclines, and taxanes. Gene signatures have the potential to transform breast cancer treatment as it becomes tailored to each patient's tumor expression profile and significantly improve the outcomes of this disease.

In the past decade, our understanding of the human genome has led to the development of many new tools to unravel the complex genetic underpinnings of a variety of diseases, including cancer. This has stimulated a large-scale effort by investigators to develop prognostic markers aimed at enabling therapeutic regimens to be tailored to specific breast cancer phenotypes, predicting sensitivity to chemotherapeutic agents and thereby avoiding overtreatment of low-risk patients and undertreatment of high-risk patients. Over the past decade, genomic technology has enabled development of a tremendous body of scientific literature describing the biologic and genetic nature of breast cancer. Such technologies, including gene expression profiling, array comparative genomic hybridization, and others have the potential to contribute greatly to our ability to 'personalize' the treatment of breast cancer.

Biomarkers are classified as either prognostic or predictive. The expression of a prognostic marker stratifies different populations of breast cancer patients with respect to the risk of an outcome, such as recurrence or disease-free survival, independent of the treatment received. Cohorts of patients involved in the development and validation of prognostic markers ideally receive uniform treatment. The goal of such studies is to identify markers that forecast the natural history of the disease and have the potential to prevent unnecessary therapy from being given to patients with a favorable prognosis, while signaling the need for more aggressive therapy for patients with a poorer prognosis. Predictive markers stratify cancer patients by response to a specific treatment regimen. Thus, cohorts of patients involved in the development and validation of predictive markers are randomized to receive the specific chemotherapy under investigation, versus a placebo or another chemotherapy regimen for comparison. Predictive markers are those that demonstrate an ability to identify a group of patients most likely to respond to a particular therapy and therefore provide guidance in the choice of therapy.

Until recently, only three individual biomarkers, estrogen receptor (ER), progesterone receptor (PGR), and human epidermal growth factor receptor-2 (HER2; also known as ERBB2 or neu), were utilized in routine clinical care to guide treatment in breast cancer patients. ER and likely PGR expression are associated with a favorable prognosis and are highly predictive of benefit from endocrine treatment.[1,2] Randomized trials have shown that tamoxifen delays recurrence and improves 10-year disease-free survival for younger and older women irrespective of nodal status.[3] Aromatase inhibitors have been demonstrated to be an effective alternative endocrine treatment in postmenopausal women.

HER2 expression is associated with more aggressive tumor behavior and a worse prognosis than that seen in patients with non-overexpressing tumors.[4-6] Its amplification is predictive of response to the HER2-targeting drug trastuzumab. While fewer than 10% of breast cancer patients benefit from the drug overall, 25–50% of patients selected on the basis of HER2 amplification are responsive.[7] Trastuzumab has been demonstrated to improve response rates, disease-free survival, and overall survival in HER2-positive patients.[8-11] In addition, HER2 status has the potential to be predictive of sensitivity to chemotherapy. Tumors that overexpress HER2 may derive greater benefit from anthracycline-based adjuvant therapy than from cyclophosphamide, methotrexate, fluorouracil (CMF)-based regimens.[12-15]

Apart from these three biomarkers, clinico-pathologic parameters are currently also integrated into prognostic stratification models, which include Adjuvant! Online (www.adjuvantonline.com), the Nottingham Prognostic Index[16], and the American Joint Committee on Cancer staging system[17] to yield outcome estimates that guide therapeutic decisions. Unfortunately, each of these tools is limited, and methods to integrate them are lacking. The result is that inaccurate risk estimation leads some patients with a good prognosis to be overtreated and experience unnecessary adverse effects, while those who are destined to fail with standard therapy cannot reliably be identified. To address this problem, prognostic and predictive tools have been developed using genomic technologies that allow for the creation of comprehensive tumor profiles as biomarkers of prognosis and sensitivity to chemotherapeutic agents.

Gene expression profiles or signatures are combinations of genes that are differentially expressed between normal and pathologic tissues or that correlate with different prognoses or phenotypes. Gene expression profiling refers to a genomic technique that measures the subset of genes that is expressed in a specific sample. The techniques of microarray and real-time reverse transcriptase PCR (RT-PCR) have been utilized to measure the expression of multiple genes simultaneously. Briefly, gene expression profiling involves quantifying gene expression and mathematically transforming these levels of expression into signatures that predict disease recurrence or treatment response. Messenger RNA (mRNA) is extracted from tumor samples and quantified by microarrays or RT-PCR to assess gene expression. Most gene profiles consist of a few to 100 genes chosen from 20 000 to 25 000 genes on a microarray or with RT-PCR. Several investigators have performed genome-gene expression analyses with oligonucleotide microarrays or commercialized chips, such as Affymetrix, whereas

© 2009 Adis Data Information BV. All rights reserved.

other groups have done focused analyses with RT-PCR on candidate genes. After gene expression is measured, hierarchical cluster analysis is performed to identify a subset of genes that show unique expression patterns for different outcomes or responses to treatment. Multigene signatures have proven to be better prognostic indicators than single biomarkers. Acharya et al.[18] combined gene expression data from genes previously identified to be associated with a poor prognosis in breast cancer to establish risk stratification. They found that expression of each individual gene did not reveal a clear distinction between the various prognostic subgroups. It was the aggregation of the genes into signatures that provided the prognostic power.

In addition to prognostic indicators, gene expression array analysis is being used to develop gene expression signatures that predict sensitivity to various chemotherapeutic agents. Tumor subtypes that are most sensitive or resistant to a given chemotherapeutic agent are identified, and gene expression data from these tumors is used to generate a signature, or gene expression profile, which is associated with a treatment response. The predictive capacity is then validated with independent tumor specimens.

Although this technology has the potential to transform the management of breast cancer, there are limitations.[19] Sample manipulation, analysis of case series selected by means of different ascertainment criteria, and cross-platform inconsistency all contribute to discordant study results. For a gene signature to become incorporated into clinical use, results acquired under a given platform must be reproducible and the sensitivity and specificity of a test should be known. Most importantly, prospective validation is critical to determine if clinical decisions influenced by the results of gene expression profiles yield more efficacious outcomes.

1. Prognostic Gene Expression Signatures in Breast Cancer

To date, seven prognostic signatures have been published, which vary by platform, degree of validation, and level of evidence supporting prognostic ability. Three have been developed into laboratory-based clinical tests that are available for widespread use: Oncotype DX® (Genomic Health, Inc., Redwood City, CA, USA), MammaPrint® (Agendia BV, Amsterdam, the Netherlands), and the *HOXB13/IL17BR* (H/I) ratio (Theros H/I^{SM}; bioTheranostics, San Diego, CA, USA). The prognostic signatures are listed in table I and described in detail below.

1.1 Molecular Subtypes Using the Intrinsic Gene Set (`Intrinsic Subtypes')

The 'intrinsic subtypes' initially consisted of four molecular classes of breast cancer (basal-like, HER2-positive, normal breast-like, and luminal/epithelial-ER-positive) differentiated by an intrinsic gene list of 496 genes[20,21] and subsequently refined with the sub-setting of luminal types A and B.[23] The subclass of basal-like tumors have low expression of ER and HER2, and amplification of genes normally expressed in the epithelium, such as cytokeratins 5, 6, and 17 (*KRT5*, *KRT6*, and *KRT17*). A second type of ER-negative tumor is the HER2-positive/ER-negative (HER2+/ER–) subtype. Basal-like and HER2+/ER– subtypes are more sensitive to chemotherapy, although these subtypes have a worse prognosis.[41] The luminal subtypes typically express HER2, as well as ER, the transcription factor GATA3, and the genes that they regulate. Luminal-A tumors express higher levels of ER and GATA3 than luminal-B tumors, and have the most favorable long-term survival after treatment with endocrine therapy.[23] Luminal-B tumors often express greater levels of HER1, HER2, and cyclin E1.

In a large-scale validation study, Carey et al. examined the intrinsic subtypes in the Carolina Breast Cancer Study and once again demonstrated significantly different overall survival ($p < 0.001$).[25] During a maximum duration of follow-up of 11.2 years with a minimum follow-up of 8.1 years, the overall survival rates were 75%, 52%, 84%, and 87% for basal-like, HER2+/ER–, luminal-A, and luminal-B subtypes, respectively. Survival was worse for the basal-like subtype (hazard ratio [HR] 1.8; 95% confidence interval [CI] 1.1, 2.9; $p = 0.03$) and the HER2+/ER– subtype (HR 3.5; 95% CI 1.9, 6.2; $p < 0.001$) than for the luminal-A subtype. The difference in survival remained when subgroups were stratified by lymph node status. This study also confirmed earlier reports that an elevated proportion of patients with *BRCA1* mutations had tumors consistent with the basal-like subtype. Another study demonstrated that HER2-/ER– tumors with cytokeratin 5 and/or 6 expression was significantly associated with *BRCA1* mutations (odds ratio [OR] = 9.0, 95% CI 1.9, 43; $p = 0.002$).[42] In addition, mutations in the *BRCA1* gene are negatively associated with HER2+ tumors.[43]

1.2 Oncotype DX® 21-Gene Profile Recurrence Score

Oncotype DX® is a 21-gene profile that was developed to estimate the risk of recurrence in newly diagnosed patients with node-negative, ER-positive, stage I or II breast cancer. The gene signature was derived from a prospectively chosen set of

© 2009 Adis Data Information BV. All rights reserved.

Table I. Breast cancer prognostic gene expression signatures

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
Intrinsic subtype predictor					
Perou et al., 2000[20]	65 tissue samples	Neoadjuvant doxorubicin (average of 16 wk)	Pretreatment open surgical biopsy; post-treatment resection; cDNA microarrays; IHC; cluster analysis	Genes with greater variation in expression between different tumors than between paired samples from the same tumor pre- and post-treatment Cluster dendrogram groupings	496 genes = 'intrinsic' gene subset Identified 4 molecular groups of mammary epithelial biology: ER+/luminal-like basal-like (ER–) HER2+ (ER–) normal breast
Sorlie et al., 2001[21]	85 tissue samples (78 carcinomas, 7 nonmalignant breast samples) Correlation with clinical outcome: 49 patients	Neoadjuvant doxorubicin, tamoxifen post-surgery	cDNA microarray (478 cDNA intrinsic clone set) hierarchical clustering analysis	Cluster dendrogram groupings Overall survival	Confirmed intrinsic subgroups found by Perou et al.[20] Luminal/ER+ subgroup can be divided into at least 2 subgroups A significant difference in overall survival ($p < 0.01$) between the subtypes (basal-like and HER2+ had the shortest survival times)
Sorlie et al., 2003[23]	122 tissue samples (including 84 samples from Sorlie et al.[21])	Neoadjuvant doxorubicin, tamoxifen post-surgery	cDNA microarrays (534 'intrinsic' genes); hierarchical clustering analysis	Cluster dendrogram groupings Time to development of distant metastases	534 'intrinsic' genes subdivided into 1 basal-like, 1 HER2+, and 2 luminal-like subgroups Time to development of distant metastasis: shortest for basal and HER2+, longest for luminal-A, and intermediate for luminal-B subtypes *BRCA1* is associated with the basal-like subtype
Carey et al., 2006[25]	496	LN– and LN+, variable treatment	IHC	Breast cancer-specific survival Follow-up: minimum duration 8.1 y; maximum duration 11.2 y	Overall survival rates: basal-like: 75% HER2+/ER–: 52% luminal-A: 84% luminal-B: 87% HRs comparing survival with the luminal-A subtype: basal-like: 1.8 (95% CI 1.1, 2.9; $p = 0.03$) HER2+/ER–: 3.5 (95% CI 1.9, 6.2; $p < 0.001$)
21-gene signature RS - Oncotype DX®					
Paik et al., 2004[26]	Training set: 477 Validation set: 668	ER+, LN– Tamoxifen No chemotherapy	RT-PCR	10-y breast cancer recurrence	21-gene signature and RS algorithm Validation set: 10-y recurrence risk: RS <18: 6.8%

Continued next page

© 2009 Adis Data Information BV. All rights reserved.

Table I. Contd

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
					RS >18 and <30: 14.3% RS >30: 30.5% HR for a 50-point change in RS = 2.8 (95% CI 1.7, 4.6)
Habel et al., 2006[27]	990 Cases: 220 (patients who died from breast cancer) Controls: 570	ER+, LN– Tamoxifen No chemotherapy	RT-PCR	10-y overall survival	Death at 10 y RS <18: 2.8% RS >18 and <30: 10.7% RS >30: 15.5%
Goldstein et al., 2008[28]	465 Cases: 99 (patients who relapsed) Controls: 366	ER+, LN–/+ Tamoxifen/aromatase inhibitor Doxorubicin-containing chemotherapy	RT-PCR	5-y recurrence	RS was a significant predictor of recurrence for LN– and LN+ patients ($p < 0.001$ for both) RS <18, 5-y recurrence: positive node: 3.3% 2–3 positive nodes: 7.9%
Albain et al., 2007[29]	361	Postmenopausal, ER+, LN+ Tamoxifen No chemotherapy	RT-PCR	10-y disease-free survival and overall survival	10-y disease-free survival: RS <18: 60% RS >18 and <30: 49% RS >30: 43% 10-y overall survival: RS <18: 77% RS >18 and <30: 68% RS >30: 51%
70-gene signature - MammaPrint®					
van't Veer et al., 2002[30]	Training set: 78 Cases: 34 (poor prognosis: patients with recurrence) Controls: 44 (good prognosis) Validation set: 19 Cases: 7 Controls: 12	ER+, LN– Mostly no receipt of tamoxifen and chemotherapy	DNA microarray (sequences selected from RefSeq)	5-y distant recurrence	70-gene signature: Predictive accuracy 65/78 (83%) (5 poor-prognosis and 8 good-prognosis patients were assigned to the opposite category by the signature) Validation set: predictive accuracy: 17/19
Van de Vijver et al., 2002[31]	295 (including 61 from training set above) Cases: 180 (poor prognosis) Controls: 115 (good prognosis)	ER+/–, LN+/– Mixed receipt of tamoxifen and chemotherapy	DNA microarray	10-y disease-free survival and overall survival	10-y disease-free survival: good prognostic signature group: 85% poor prognostic signature group: 51% (HR = 5.1 comparing distant metastases among poor and good prognostic signature groups) 10-y overall survival: good prognostic signature group: 95% poor prognostic signature group: 55%

Continued next page

© 2009 Adis Data Information BV. All rights reserved.

Table I. Contd

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
Buyse et al., 2006[32]	302 Adjuvant! Online/ MammaPrint® low/low risk: 52 low/high risk: 28 high/low risk: 59 high/high risk: 163	ER+/–, LN– No receipt of tamoxifen or chemotherapy	DNA microarray	5- and 10-y distant metastases and 10-y overall survival	HR for time to distant metastases: MammaPrint®: 2.32 Adjuvant! Online: 1.68 HR for overall survival: MammaPrint®: 2.79 Adjuvant! Online: 1.67 MammaPrint® 10-y overall survival: high-risk group: 0.69 (regardless of Adjuvant! Online group) low-risk group: 0.88 and 0.89, for low- and high-risk Adjuvant! Online groups, respectively
The H/I test (*HOXB13*/*IL17BR* ratio)					
Ma et al., 2004[33]	Training set: 60 Cases: 28 (patients with recurrences) Controls: 32 Validation set: 20 Cases: 10 Controls: 10	ER+, LN+/– Tamoxifen No chemotherapy	Microarray analysis; RT-PCR	Disease-free survival	*HOXB13*/*IL17BR* ratio Validation set: predictive accuracy: 16/20
Reid et al., 2005[34]	58	ER+, LN+/– Tamoxifen No chemotherapy	RT-PCR	Disease-free survival	Did not confirm prognostic signature: OR = 1.30 (95% CI 0.88, 1.93; p = 0.18) for recurrence risk
Goetz el al., 2006[24]	206 (LN–: 130)	ER+, LN+/– Tamoxifen No chemotherapy	RT-PCR	Relapse-free, disease-free, and overall survival	HR RFS: 1.5 (95% CI 0.93, 2.27), disease-free survival: 1.6 (95% CI 1.04, 2.38), overall survival: 1.29 (95% CI 0.81, 2.08) when limited to LN– patients
Ma et al., 2006[22]	852 (ER+, LN–, tamoxifen-treated and untreated: 225)	ER+/–, LN+/–, Tamoxifen-treated and untreated	RT-PCR	5-y RFS	Expression of *HOXB13* was associated with shorter RFS (p = 0.008) and expression of *IL17BR* was associated with longer RFS (p < 0.0001) Significant predictor of RFS, HR = 3.9; 95% CI 1.5, 10.3; p = 0.007 for ER+, LN– subgroup
Jansen et al., 2007[35]	468	ER+, LN–, No tamoxifen No chemotherapy	RT-PCR	Disease-free survival, overall survival (median of 6 y follow-up)	HR disease-free survival: 1.06 (95% CI 1.02, 1.10; p = 0.001) overall survival: 1.07 (95% CI 1.03, 1.10; p < 0.001), when the ratio is analyzed as a univariate continuous variable

Continued next page

© 2009 Adis Data Information BV. All rights reserved.

Table I. Contd

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
Rotterdam signature					
Wang et al., 2005[36]	Training set: 115 Validation set: 171	LN–, ER+/–, No systemic treatment	Banked frozen tumor samples; Affymetrix Human U133a GeneChips	Distant metastases within 5 y	78-gene signature Validation set: prediction of 5-y metastatic disease: sensitivity: 93% specificity: 48% Multivariate regression estimation of HR for the occurrence of tumor metastasis within 5-y: 5.55 ($p < 0.0001$)
Foekens et al., 2006[37]	180	LN–, ER+/–, No systemic treatment	Frozen tumor samples; custom-designed DNA chips	5- and 10-y distant metastasis-free survival	5-y distant metastasis-free survival: GPS group: 96% PPS group: 74% sensitivity: 90% specificity: 50% PPV: 38% NPV: 94% 10-y distant metastasis-free survival: GPS group: 94% PPS group: 65%
Desmedt et al., 2007[38]	198	LN–, ER+/–, No systemic treatment	Frozen tumor samples; Affymetrix Human U133a GeneChips	5- and 10-y distant metastasis-free survival 5- and 10-y overall survival	5-y distant metastasis-free survival: GPS group: 98% PPS group: 76% 10-y distant metastasis-free survival: GPS group: 94% PPS group: 73% 5-y overall survival: GPS group: 98% PPS group: 84% 10-y overall survival: GPS group: 87% PPS group: 72%
Wound response indicator					
Chang et al., 2005[39]	Training set and validations set: 295	LN+/–, variable treatment	Stanford cDNA microarrays and Rosetta/NKI oligonucleotide microarrays; cluster linkage of 'core serum response' genes	First bifurcation of the clustering dendrogram identified an 'activated' vs 'quiescent' wound response signature 10-y overall survival and distant	512-gene signature Validation study: activated: n = 126/295 quiescent: n = 169/295 10-y overall survival: activated wound response signature: 50% quiescent wound response signature: 84% 10-y distant metastasis-free probability:

Continued next page

© 2009 Adis Data Information BV. All rights reserved.

Table I. Contd

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
				metastasis-free probability	activated wound response signature: 51% quiescent wound response signature: 75%
Invasive gene signature					
Liu et al., 2007[40]	Training set: 9 (derivation of 186 genes) Cases: 6 Controls: 3 (normal breast tissue) 295 (Netherlands Cancer Institute database) 286 (Erasmus Medical Center database) Validation set: 284	ER+/–, LN+/–, –	Comparison of the gene-expression profile of CD44+CD24–/low tumorigenic breast cancer cells with normal breast epithelium; microarrays; RT-PCR; cluster linkage	10-y metastasis-free survival and overall survival	10-y disease-free survival and overall survival was predicted by the IGS ($p < 0.001$ for both) Validation study: IGS combined with NIH prognostic criteria GPS (negative IGS correlation): n = 60/284 PPS (positive IGS correlation): n = 224/284 10-y metastasis-free survival: GPS: 81% PPS: 57%

cDNA = complementary DNA; **ER** = estrogen receptor; **GPS** = good prognostic signature; **HER2** = human epidermal growth factor receptor-2; **HR** = hazard ratio; **IGS** = 'invasiveness' gene signature; **IHC** = immunohistochemistry; **LN** = lymph node; **NIH** = US National Institutes of Health; **NKI** = Netherlands Cancer Institute; **NPV** = negative predictive value; **OR** = odds ratio; **PPS** = poor prognostic signature; **PPV** = positive predictive value; **RFS** = relapse-free survival; **RS** = recurrence score; **RT-PCR** = reverse transcriptase PCR.

250 candidate genes, which were measured in 447 patients with breast cancer obtained from three preliminary studies.[44-46] From the 250 genes, 21 genes (16 cancer-related genes and 5 reference genes) best predicted 10-year breast cancer recurrence. The cancer-related genes include a proliferation group (Ki-67 [*MKI67*], STK15 [*AURKA*], survivin [*BIRC5*], cyclin B1 [*CCNB1*], *MYBL2*), *HER2* and its coregulated gene *GRB7*, estrogen-related genes (*ER*, *PGR*, *BCL2*, and *SCUBE2*), a recurrence group (beta-actin [*ACTB*], *GAPDH*, *RPLP0*, *GUS*, and *TFRC*), invasion genes (stromelysin 3/matrix metalloproteinase 11 [*MMP11*] and cathepsin L2 [*CTSL2*]) and *GSTM1*, *CD68*, and *BAG1*. Expression levels of these genes were measured by RT-PCR and then placed in a quantitative algorithm to produce the recurrence score (RS), a number between 0 and 100. The RS is correlated with a continuous measure of recurrence risk, though three distinct risk categories have been developed: low (RS <18), intermediate (RS >18 but <30), or high (RS >30). To test the clinical validity of the 21-gene signature, Paik et al.[26] generated recurrence scores for 668 women who had participated in a randomized, controlled trial conducted by the National Surgical Adjuvant Breast and Bowel Project (NSABP-B14) who were treated with tamoxifen. Stratification into the three risk categories yielded 10-year recurrence risks of 6.8%, 14.3%, and 30.5% for the low-, intermediate-, and high-risk groups, respectively ($p < 0.001$). A multivariable analysis demonstrated that the RS was the strongest predictor of recurrence, independent of traditional risk factors (HR 2.8; 95% CI 1.7, 4.6) for a 50-point change in the RS. In a second validation study, Habel et al.[27] examined the predictive value of the RS in a community-based case-control study of 875 lymph node-negative, ER-positive patients, of whom 205 were treated with tamoxifen. The probability of death at 10 years was 2.8%, 10.7%, and 15.5% in the low-, intermediate-, and high-risk groups, respectively. The prognostic value of the RS persisted after adjustment for tumor grade and disease stage.

Although the 21-gene signature was developed for outcome prognostication in node-negative patients, two studies have provided validation for the use of this assay in node-positive patients. Goldstein et al.[28] evaluated the prognostic utility in either node-negative or node-positive, ER-positive patients treated with doxorubicin-containing chemotherapy. 465 patients with 0–3 positive axillary nodes who had participated in an Eastern Cooperative Oncology Group study of doxorubicin/cyclophosphamide versus doxorubicin/docetaxel were included in this biomarker substudy. The RS was a highly significant predictor of recurrence, including node-negative and node-positive disease ($p < 0.001$ for both) and when adjusted for

© 2009 Adis Data Information BV. All rights reserved.

other clinical variables. The risk of recurrence was high in patients with 2–3 positive nodes, compared with those with 0–1 positive node. For an RS value of <18, 3.3% (95% CI 2.2, 5.0) of patients with 0–1 positive node experienced a recurrence in 5 years, versus 7.9% (95% CI 4.3, 14.1) of patients with 2–3 positive nodes. The RS also predicted recurrence more accurately than clinico-pathologic variables when integrated by the Adjuvant! Online model adjusted for 5-year rather than 10-year outcomes. The study demonstrated that even node-positive patients may be exempt from standard aggressive chemotherapy regimens on the basis of a low RS. Albain et al.[29] further validated the use of the 21-gene signature as a prognostic indicator of 10-year disease-free survival and overall survival in node-positive, ER-positive patients.

1.3 MammaPrint® 70-Gene Profile

MammaPrint® is a 70-gene prognostic signature developed by van't Veer et al. from the Netherlands Cancer Institute. The MammaPrint® assay requires fresh-frozen tissue for analysis. The initial profile was developed using 78 lymph node-negative patients younger than 55 years who had tumors that were less than 5 cm in diameter and who did not carry a breast cancer gene mutation, with the goal of predicting the 5-year distant recurrence risk.[30] Patients were separated into two groups based on disease-free survival after 5 years from diagnosis. The gene expression profiles of the two groups were compared to derive a 70-gene profile that could predict the clinical outcome. These authors subsequently validated the 70-gene signature on 295 patients (including 61 patients from the original study) younger than 53 years with stage I or II breast cancer and tumors less than 5 cm in diameter.[31] Unlike the initial training set, the validation population included patients with both negative and positive lymph node disease and ER status, and was mixed in terms of receipt of chemotherapy and tamoxifen. At 10 years, patients with a 'good-prognosis' 70-gene signature were more likely to remain free of distant metastases (85% vs 51%) and achieved better overall survival (95% vs 55%) than patients with a 'poor-prognosis' signature (HR 5.1; 95% CI 2.9, 9.0; $p < 0.001$). The HR continued to remain significant when the study cohort was separately analyzed based on node status. Results excluding the 61 patients from the training set were similar. Further analysis demonstrated that the 70-gene signature reclassified patients when compared with previously used classification tools, including the St Gallen criteria[47] and US National Institutes of Health (NIH) risk stratification algorithms. Many patients with the good-prognosis signature were reassigned to the low-risk group after evaluation with traditional methods. The 70-gene signature placed 40% of patients from the study by van de Vijver et al.[31] into the good-prognosis group while the St Gallen index and the NIH criteria placed 15% and 7% into the good-prognosis group, respectively. The 70-gene signature was found to more accurately risk stratify patients for a given outcome than the St Gallen index and the NIH criteria. Low-risk patients identified by the gene signature had a lower rate of distant disease recurrence than those classified as low risk by the St Gallen or NIH criteria, and high-risk patients identified by the signature were more likely to develop metastatic disease than high-risk patients classified by the traditional methods.

The TransBIG research network (http://www.breastinternationalgroup.org/TRANSBIG) also performed a validation of the MammaPrint® assay, which involved 302 women younger than 61 years with lymph node-negative, stage I and II disease not treated with chemotherapy or tamoxifen.[32] In this study, the 70-gene signature yielded independent prognostic information beyond the Adjuvant! Online tool, and by comparison with clinico-pathologic information, more accurately predicted distant metastases (HR = 2.32; 95% CI 1.35, 4.00 vs HR = 1.68; 95% CI 0.92, 3.07) and overall survival (HR = 2.79; 95% CI 1.60, 4.87 vs HR 1.67; 95% CI 0.93, 2.98). High-risk patients identified by MammaPrint® had a 10-year overall survival of 0.69 regardless of the risk group assigned by Adjuvant! Online criteria. For patients assigned to the low-risk group by MammaPrint®, the 10-year survival rates were 0.88 and 0.89 for those in the low- and high-risk Adjuvant! Online groups, respectively.

1.4 The *HOXB13/IL17BR* Ratio (H/I Test)

The Theros H/I^{SM} test is based on a 2-gene signature (*HOXB13* and *IL17B*) developed by Ma et al.[33] for use in paraffin-embedded tissues. The authors identified two genes in 60 patients with ER-positive, lymph node-positive or lymph node-negative breast cancer treated with tamoxifen that were highly associated with outcome. High expression of *HOXB13* predicted recurrence, and high expression of *IL17BR* predicted non-recurrence. A higher ratio of the two genes strongly predicted recurrence in this training set. Ectopic expression of *HOXB13* in MCF10A breast endothelial cells enhances motility and invasion *in vitro*, and expression is increased in both preinvasive and invasive breast cancer.

Validation studies helped revise the assay method until prognostic accuracy was optimized. Initially, Reid et al.[34] found no relationship between the ratio and distant relapse in a cohort of 58 patients with resectable, ER-positive breast cancer.

© 2009 Adis Data Information BV. All rights reserved.

Ma et al.[22] performed a further validation trial on 852 tumors in patients with stage I or II breast cancer with a median follow-up of 6.8 years. A revised method was used to combine and normalize the expression of the two genes into an index that is now the basis of the H/I assay. The ratio was predictive only in patients with node-negative, ER-positive disease. The adjusted HR incorporating other risk factors was 3.9 (95% CI 1.5, 10.3; p = 0.007) regardless of tamoxifen treatment. Jansen et al. further evaluated the two-gene ratio in 468 ER-positive, node-negative, tamoxifen-untreated patients and found that the ratio was associated with poor disease-free survival (HR = 1.06; 95% CI 1.02, 1.10; p = 0.001) and poor overall survival (HR = 1.07; 95% CI 1.03, 1.10; p < 0.001) when the ratio was analyzed as a univariate continuous variable.[35] Goetz et al.[24] subsequently evaluated 206 ER-positive patients treated only with tamoxifen in a North Central Cancer Treatment Group clinical trial. Expression values were normalized by a different method than that used by Ma et al.,[22] and a different cutoff point was calculated for the ratio. The ratio had modest predictive strength for relapse-free (HR 1.5; 95% CI 0.93, 2.27), disease-free (HR 1.6; 95% CI 1.04, 2.38), and overall survival (HR 1.29; 95% CI 0.81, 2.08) when limited to node-negative patients.

1.5 Rotterdam Signature

The Rotterdam Signature is a 76-gene microarray assay that was developed in lymph node-negative patients, regardless of age, tumor size, grade, and hormone receptor status, and aimed to predict metastatic disease within 5 years.[36] 286 patients composed both the training set (n = 115) and the initial validation group (n = 171). The validation study yielded sensitivity for prediction of metastatic recurrence of 93% and specificity of 48% (HR 5.67 uncorrected for conventional prognostic factors and HR 5.55 corrected for these factors). This group performed a multicenter validation study of 180 patients with stage I and II breast cancer, showing 5- and 10-year distant metastasis-free survival rates of 96% and 94%, respectively, for the good-prognosis group and 5- and 10-year distant metastasis-free survival rates of 74% and 65%, respectively, for the poor-prognosis group.[37] A further validation study including 198 lymph node-negative patients resulted in 5- and 10-year distant metastasis-free survival rates of 98% and 94% for the good-prognosis group and 76% and 73% for the poor-prognosis group.[38] This signature overlaps with only three genes that compose the 70-gene assay of MammaPrint®. Like MammaPrint®, it requires whole sections of frozen tissue instead of core biopsy samples.

1.6 Wound Response Indicator

The wound response indicator (WRI) expression signature was derived from the transcriptional response of normal fibroblasts to serum in cell culture.[48] The concept that distant metastases are more likely among patients whose breast cancers have activated pathways for matrix remodeling, cell motility, and angiogenesis than among those that do not is the rationale underlying the development of this signature. In a validation study, 295 patients with early-stage breast cancer revealed that patients whose tumors expressed the WRI had significantly shorter overall survival and distant metastasis-free survival than patients whose tumors did not express the signature.

1.7 Invasive Gene Signature

It is theorized that only a small portion of cells within a tumor are tumorigenic. In breast cancer, a small subgroup of cells have low or undetectable CD24 expression and high CD44 expression, and these cells have been demonstrated to generate tumors when injected into immunocompromised mice.[49] The majority of cells in a cancer are nontumorigenic and are unable to give rise to new tumor. The CD44+/CD24– cells are hypothesized to be 'breast cancer stem cells.' A 186-gene profile, termed the invasive gene signature (IGS), was derived by comparing normal breast epithelial cells with CD44+/CD24– breast cells.[50] The expression of genes in the IGS is associated with shorter overall survival and metastasis-free survival times (p < 0.001). When combined with NIH prognostic criteria, the IGS predicted a 10-year, metastasis-free survival rate of 81% among patients in the good-prognosis group and 57% among patients in the poor-prognosis group.

1.8 Comparison of Prognostic Signatures

The methods of the signatures differ in that Oncotype DX® and H/I are done in formalin-fixed, paraffin-embedded tumor tissues and utilize RT-PCR, while fresh unfixed tumor tissue is required for the DNA microarrays used in MammaPrint®. The commercially available gene signatures have minimal overlap; the 21-gene profile and the 70-gene profile share only one gene in common. Fan et al. evaluated the predictive agreement between several profiles, including Oncotype DX®, H/I, MammaPrint®, Intrinsic Subtype, and Wound Response.[51] Oncotype DX® RS, the H/I ratio, Intrinsic Subtype, and Wound Response were estimated from microarray data on the same 295 samples that had been used to develop the 70-gene signature. The 70-gene signature and the derived RSs, as well as

© 2009 Adis Data Information BV. All rights reserved.

the intrinsic subtype profiles and wound response signatures, predicted overall survival and disease-free survival, but the ratio did not predict either outcome. Most tumors associated with a poor prognosis based on an intrinsic subtype, such as basal-like, HER+/ER– or luminal B, were also classified as having a poor-prognosis 70-gene profile, activated wound response, and high RS. The two best validated models, Oncotype DX® and MammaPrint®, were specifically compared in the same cohort of 295 patients involved in the development of the 70-gene signature. Thus, the comparison was expected to favor the 70-gene signature. The intermediate- and high-risk RS groups were combined and compared with the poor-prognosis group of the 70-gene signature. The agreement between Oncotype DX® and MammaPrint® was 81% (239 of 295 patients). Despite minimal gene overlap, both Oncotype DX® and MammaPrint® were found to be precise prognostic indicators.

2. Molecular Profiles Predicting Responsiveness to Chemotherapy

Predictive signatures indicate how patients will respond to a given course of adjuvant treatment. As prognostic signatures were being developed, the logical outgrowth of this development was to determine whether they would also be predictive of chemotherapy benefit. For example, such a signature could be used to determine whether a patient had a sufficiently low risk of both recurrence *and* benefit from chemotherapy, indicating that such therapy should not be given. This could also aid in determining which patients at low risk by standard prognostic measures would benefit from the addition of chemotherapy. To date, the most data have been accumulated on the Oncotype DX® profile in this regard, though similar studies are under way with other signatures (table II).

2.1 Oncotype DX®

Several studies have evaluated the utility of Oncotype DX® in determining which breast cancer patients receive additional benefit from systemic chemotherapy. Paik et al.[52] studied a subset of 651 patients with ER-positive, lymph node-negative disease who were randomly assigned to receive either tamoxifen alone or tamoxifen with (C)MF (methotrexate and fluorouracil with or without cyclophosphamide) in the NSABP B-20 trial. In this trial, an overall benefit in disease-free survival was seen with the addition of chemotherapy to tamoxifen. The risk of distant recurrence was compared between the tamoxifen-alone arm and the tamoxifen with chemotherapy arm to determine whether there was a benefit with the addition of chemotherapy when stratified by RS. When the data were stratified by risk group, a significant benefit was restricted to patients with a high RS (RS ≥31). Patients with a high RS had a significantly lower overall risk of recurrence (relative risk [RR] 0.26; 95% CI 0.13, 0.53) and an absolute decrease in the 10-year distant recurrence rate (mean 27.6%; standard error [SE] 8.0%). Patients with a low RS (<18) did not appear to receive any benefit from chemotherapy treatment (RR 1.13; 95% CI 0.46, 3.78; absolute decrease in distant recurrence rate at 10 years: mean −1.1%; SE 2.2%). The group of patients with intermediate-RS tumors was too small to draw a definitive conclusion, with a very wide CI (HR 0.61; 95% CI 0.24, 1.59). Overall, this retrospective study suggested that ER-positive, lymph node-negative patients with a high RS would benefit from chemotherapy, while patients with a good prognosis, treated with tamoxifen, may be able to be spared chemotherapy. Because this study could not definitely determine whether patients with an intermediate-risk RS would benefit from additional systemic therapy, an Intergroup Trial, led by the Eastern Cooperative Oncology Group, was launched. The Trial Assigning Individualized Options for Treatment (TAILORx) compares disease-free survival among women with node-negative disease who have an RS between 11 and 25 and are randomized to receive either adjuvant chemotherapy plus tamoxifen or tamoxifen alone.[57]

Response to chemotherapy can be indicated by the pathologic response of the definitive surgical specimen following neoadjuvant chemotherapy, and a pathologic complete response (pCR) is a surrogate measure of long-term, disease-free survival. Two studies have examined whether the RS predicted pathologic response in patients who received neoadjuvant chemotherapy. Gianni et al. found that the RS predicted complete response after treatment with a combination of anthracycline and taxane therapy,[53] while Mina et al. found no such relationship.[50] Chang et al. assessed chemotherapy response prediction in 12 patients (of 72 in total) with a complete clinical response in a docetaxel trial and found that a high RS was associated with a complete response ($p = 0.008$).[54] When the RS was used as a continuous variable, a 14-unit increase (the difference between the high- and low-risk groups) was modestly predictive of a clinical complete response (OR 1.7; CI 1.15, 2.60).

Albain et al.[29] evaluated the predictive accuracy of Oncotype DX® in stratifying patients based on potential benefit from chemotherapy in a subset of node-positive patients who were enrolled in a study conducted by the Southwest Oncology

© 2009 Adis Data Information BV. All rights reserved.

Table II. Breast cancer predictive gene expression signatures

Study, year	Patients analyzed (no.)	Patient characteristics	Method	Endpoints	Results
21-gene signature RS-Oncotype DX®					
Paik et al., 2006[52]	651 Tamoxifen-only subgroup: 227 Tamoxifen with chemotherapy subgroup: 424	ER+, LN– Randomly assigned to receive tamoxifen or tamoxifen with [C]MF	RT-PCR	10-y disease-free survival or distant recurrence	RR of recurrence: high RS: 0.26 (95% CI 0.13, 0.53) low RS: 1.13 (95% CI 0.46, 3.78) Patients with high RS benefited from [C]MF, while patients with low RS did not
Gianni et al., 2005[53]	89	Doxorubicin and paclitaxel	RT-PCR	pCR in breast tissue and lymph nodes	RS was positively associated with the likelihood of pCR ($p = 0.005$)
Mina et al., 2007[50]	45	Doxorubicin every 2 wk × 3 cycles and docetaxel weekly × 6 cycles	Core biopsies RT-PCR	pCR: 6/45 (14%)	RS did not correlate with pCR ($p = 0.67$)
Chang et al., 2008[54]	72	Docetaxel every 3 wk × 4 cycles	Core biopsies RT-PCR	Complete clinical response: 12/72 (17%)	Complete response was more likely with a high RS ($p = 0.008$)
Albain et al., 2007[29]	367 Tamoxifen-only subgroup: 148 Tamoxifen with chemotherapy subgroup: 219	ER+, LN+ Randomly assigned to receive tamoxifen or tamoxifen with CAF	RT-PCR	10-y disease-free survival	10-y disease-free survival: high RS: significant benefit among chemotherapy arm ($p = 0.003$) low RS: no benefit ($p = 0.97$)
The H/I test (*HOXB13*/*IL17BR* ratio)					
Jerevall et al., 2008[55]	264 (ER-positive patients: 179)	Postmenopausal, majority ER+, LN+ Randomized to receive 2 or 5 y of tamoxifen therapy	RT-PCR	Recurrence-free survival	Recurrence rate ratio for ER+ patients with low ratio comparing 5 y and 2 y of tamoxifen: RR = 0.39 (95% CI 0.17, 0.91; $p = 0.030$)
Goetz et al., 2008[56]	110	ER+, LN– Randomized to receive 5 y of tamoxifen	*CYP2D6* genotyping and assessment of prescribed inhibitors RT-PCR for *HOXB13*/*IL17BR* expression (H/I ratio)	Disease-free survival and overall survival	Combined *CYP2D6* and H/I risk factor is significantly associated with disease-free survival ($p = 0.004$) and overall survival ($p = 0.009$) HR for overall survival relative to high *CYP2D6*, low H/I ratio: either low *CYP2D6* or high ratio: 2.41 (95% CI 1.08, 5.37; $p = 0.031$) low *CYP2D6* and high H/I ratio: 3.15 (95% CI 1.17, 8.52; $0 = 0.024$)

CAF = cyclophosphamide, doxorubicin, fluorouracil; **[C]MF** = methotrexate and fluorouracil with or without cyclophosphamide; **CYP** = cytochrome P450; **ER** = estrogen receptor; **HR** = hazard ratio; **LN** = lymph node; **pCR** = pathologic complete response; **RR** = relative risk; **RS** = recurrence score; **RT-PCR** = reverse transcriptase PCR.

Group, evaluating disease-free survival with cyclophosphamide, doxorubicin, and fluorouracil (CAF) chemotherapy compared with tamoxifen alone. They showed that there was no benefit in 10-year disease-free survival in patients with a low RS (<18) from treatment with CAF chemotherapy compared with tamoxifen alone ($p = 0.97$). In contrast, among patients with a high RS (>31), a significant improvement in 10-year, disease-free survival was shown in the group that received CAF ($p = 0.033$); the effects on overall survival were not reported. Thus, ER-positive patients with a low RS

© 2009 Adis Data Information BV. All rights reserved.

(regardless of node status) appeared to gain benefit from systemic chemotherapy in addition to tamoxifen. Although guidelines recommend chemotherapy for all node-positive breast cancer patients, the study by Albain et al.[20] suggests that there is a group of node-positive patients who may be overtreated and are incurring unnecessary adverse effects from chemotherapy.

2.2 MammaPrint®

The MINDACT (Microarray in Node Negative Disease May Avoid Chemotherapy) trial is currently evaluating the clinical utility of MammaPrint® in selecting lymph node-negative breast cancer patients to receive adjuvant chemotherapy through comparison with Adjuvant! Online. All patients will have risk of relapse assessed by both MammaPrint® and Adjuvant! Online. If both prognostic indicators assign the patient to a low-risk group or a high-risk group, then chemotherapy will be withheld or administered, respectively. However, if the methods yield discordant results, then patients will be randomly assigned to MammaPrint® or Adjuvant! Online to determine treatment. The primary goal of this study is to determine whether a significant proportion of patients who would normally receive chemotherapy based on clinicopathologic factors will be spared chemotherapy without negatively affecting their survival.

2.3 The H/I Test

Trials on the ability to select high-risk patients for chemotherapy on the basis of the *HOXB13*/*IB17BR* ratio are necessary for predictive validation. The ratio has, to date, been tested as a predictor of tamoxifen benefit. Jerevall et al. assessed the benefit of hormone therapy among patients with ER-positive tumors randomized to 2 years versus 5 years of therapy.[55] They found that among patients with a low ratio (or low expression of *HOXB13* alone), 5 years of tamoxifen treatment provided significant benefit compared with 2 years of treatment in terms of disease-free survival, leading them to conclude that high *HOXB13* expression may be correlated with tamoxifen resistance. Goetz et al.[56] hypothesized that a combined cytochrome P450 (CYP) 2D6 (*CYP2D6*) and *HOXB13*/*IB17BR* index would be a better indicator of tamoxifen response, given that the *CYP2D6**4 genotype modifies tamoxifen metabolism. Node-negative, ER-positive breast cancer patients enrolled in a tamoxifen-only arm of the North Central Cancer Treatment Group study 89-30-52 were evaluated. The study showed not only that the combination of CYP2D6 metabolism and *HOXB13*/*IB17BR* variation influences the risk of recurrence and survival, but also that the combined index was further predictive of response to tamoxifen therapy.

3. Profiles Predicting Response to Specific Chemotherapeutic Agents

In addition to a global signature to predict need for or response to chemotherapy or endocrine therapy in a nonspecific manner, there have also been a variety of efforts to develop gene signatures that predict response to specific chemotherapy agents or regimens. These signatures and their details are listed in table III (for anthracycline-based regimens) and table IV (for taxane-based regimens). The majority of signatures predict sensitivity to anthracycline-based regimens. Ayers et al. developed a 74-gene signature that predicted pCR with 78% accuracy in patients who underwent treatment with paclitaxel followed by fluorouracil, doxorubicin, and cyclophosphamide (FAC).[58] The 86-gene profile developed by Gianni et al. was predictive of pCR in patients who received a combination of anthracycline and taxane therapy.[53] Furthermore, Cleator et al.[61] were able to stratify patients as doxorubicin, cyclophosphamide (AC) sensitive or resistant ($p = 0.04$) based on their 253-gene signature. Dressman et al.[62] were unable to develop a predictive signature for pCR in patients treated with liposomal doxorubicin and paclitaxel combined with local whole breast hyperthermia. However, they formed a 33-gene signature that was predictive of persistent lymph node involvement and a 22-gene signature predictive of the inflammatory breast cancer phenotype. Modlich et al.[64] built a 59-gene signature predictive for pathologic responses, as well as a 31-gene profile for favorable outcomes and a 26-gene profile for poor outcomes in patients receiving epirubicin and cyclophosphamide. Yang et al.[66] showed that gene ontology classes for vascular endothelial growth factor receptor (VEGFR) activity and mitosis activity were associated with a response to an anthracycline-based regimen combined with bevacizumab.

Two studies attempted to validate the predictive capacity of the Intrinsic Gene Set. Rouzier et al.[59] showed that the gene profiles defined by the subtypes that comprise the Intrinsic Gene Set are predictive of pCR in patients who received paclitaxel followed by FAC. They concluded that basal-like and HER2+ subtypes are more sensitive to paclitaxel and doxorubicin than the luminal subtype. Carey et al.[41] also evaluated the predictive potential of the Intrinsic Gene Set in patients treated with AC or an AC-taxane regimen. The clinical complete response and pCR was higher for the basal-like and HER2+/ER− subtypes than for the luminal subtype

© 2009 Adis Data Information BV. All rights reserved.

© 2009 Adis Data Information BV. All rights reserved.

Table III. Predictive gene signatures of response to anthracycline-based regimens

Study, year	Patients analyzed (no.)	Treatment	Method	Endpoints	Results
Ayers et al., 2004[58]	Training set: 24 Validation set: 18	Paclitaxel weekly ×12 cycles followed by FAC	FNA, flash frozen (RNA later); cDNA arrays	pCR Training set: 6/24 Validation set: 7/18	74-gene signature: predictive accuracy: 78% sensitivity: 43% specificity: 100% PPV: 100% NPV: 73%
Rouzier et al., 2005[59]	82	Paclitaxel weekly ×12 cycles followed by FAC	FNA; Affymetrix U133A chip	pCR 21/62 (26%)	'Intrinsic gene set' pCR rates per subtype: basal-like 45% ErbB2 45% luminal 6% normal-like 0%
Gianni et al., 2005[53]	Training set: 89 Validation set: 82	Training set: doxorubicin and paclitaxel every 3 wk × 3 cycles, followed by paclitaxel weekly ×12 cycles, postoperative CMF Validation set: paclitaxel weekly × 12 cycles, followed by FAC ×4 cycles	Training set: Core biopsy; paraffin sections; RT-PCR Validation set: FNA; Affymetrix chip	pCR in breast tissue Training set: 11/89 (12%); pCR in breast tissue and lymph nodes Validation set: 21/82 (26%)	86-gene signature and RS from 21-gene profile was predictive of pCR Validation set: 24 genes ($p < 0.05$) 32 genes ($p < 0.1$)
Hannemann et al., 2005[60]	48	AC ×6 cycles vs AD ×6 cycles	Core biopsy, snap frozen; cDNA	'Near' pCR 10/48 (20%)	No predictive signature
Cleator et al., 2006[61]	40	AC ×6 cycles	Core biopsy; Affymetrix U133A chip	Clinical complete response: 22/40 Partial response: 7/40 Stable disease: 11/40	253-gene signature predictive accuracy: 67%
Dressman et al., 2006[62]	37	Liposomal doxorubicin, paclitaxel combined with local whole breast hyperthermia every 3 wk	Ultrasound-guided core biopsy; Affymetrix U133 Plus 2.0	pCR 3/37 Persistent lymph node involvement 27/37 IBC phenotype 14/37	No predictive gene signature for pCR 33-gene signature for persistent lymph node involvement 22-gene signature for IBC phenotype
Carey et al., 2007[41]	107 basal-like: 34 HER2+/ER–: 11 luminal: 62	AC or AC-taxane ×4 cycles	UNC Lineberger Comprehensive Cancer Center Neoadjuvant Database; IHC for molecular subtyping	Clinical response pCR	Clinical response: basal-like: 85% HER2+/ER-: 70% luminal: 47% pCR: basal-like: 27% HER2+/ER-: 36% luminal: 7%

Continued next page

Table III. Contd

Study, year	Patients analyzed (no.)	Treatment	Method	Endpoints	Results
Bonnefoi et al., 2007[63]	125	Non-taxane regimen (FEC × 6 cycles) vs taxane regimen (TET × 3 cycles)	Frozen biopsies; Affymetrix X3P microarrays	pCR Total: 55/125 FEC group: 28/66 TET group: 27/59	Regimen-specific signatures (produced from combined single-agent sensitivity signatures) FEC predictor: sensitivity: 96% specificity: 66% PPV: 68% NPV: 96% TET predictor: sensitivity: 93% specificity: 69% PPV: 71% NPV: 92%
Modlich et al., 2005[64]	Training set: 56 Validation set: 27	Epirubicin and cyclophosphamide every 2 wk × 4 cycles	Needle biopsies; Affymetrix HG-U133A; RT-PCR	pCR Training set: 8/56 Validation set: 4/27 Partial remission Training set: 40/56 Validation set: 19/27 No change Training set: 8/56 Validation set: 4/27	59-gene signature Specificity pCR: >74% Partial remission: 100% No change: >62%
Salter et al., 2008[65]	133	TFAC	Affymetrix U13A	Responders (pCR) 33/133 Nonresponders (non-pCR) 99/133	Combined TFAC signature (from NCI-60 individual chemotherapy sensitivity signatures): sensitivity: 59% specificity: 63% NPV: 82% Sensitivity to chemotherapy was higher in responders than in nonresponders ($p = 0.002$)
Yang et al., 2008[66]	21 (20 inflammatory breast cancer)	Bevacizumab × 1 cycle followed by bevacizumab plus doxorubicin and docetaxel every 3 wk × 6 cycles	Core biopsy; IHC; Agilent whole human genome arrays	Partial response: 14/21 (67%) Stable disease: 5/21 Progressive disease: 2/21	Gene ontology classes for VEGFR activity and mitosis activity were associated with a clinical response (partial response)

AC = doxorubicin, cyclophosphamide; **AD** = doxorubicin, docetaxel; **cDNA** = complementary DNA; **CMF** = cyclophosphamide, methotrexate, fluorouracil; **ER** = estrogen receptor; **FAC** = fluorouracil, doxorubicin, cyclophosphamide; **FEC** = fluorouracil, epirubicin, cyclophosphamide; **FNA** = fine needle aspirate; **HER2** = human epidermal growth factor receptor-2; **IBC** = inflammatory breast cancer; **IHC** = immunohistochemistry; **NPV** = negative predictive value; **pCR** = pathologic complete response (no pathologic evidence of invasive breast cancer); **PPV** = positive predictive value; **RS** = recurrence score; **RT-PCR** = reverse transcriptase PCR; **TET** = docetaxel × 3 cycles followed by epirubicin plus docetaxel; **TFAC** = paclitaxel, fluorouracil, adriamycin [doxorubicin], cyclophosphamide; **VEGFR** = vascular endothelial growth factor.

© 2009 Adis Data Information BV. All rights reserved.

Table IV. Predictive gene signatures of response to taxane-only regimens

Study, year	Patients analyzed (no.)	Treatment	Method	Endpoints	Results
Chang et al., 2003,[67] 2005[48]	Training set: 24 Validation set: 6	Docetaxel every 3 wk × 4 cycles	Core biopsies; Affymetrix HgU95-Av2 gene chip	Chemotherapy sensitive (<25% residual tumor) 11/24 Chemotherapy resistant (>25% residual tumor) 13/24	92-gene signature (includes the 21-gene signature): sensitivity: 85% specificity: 90% PPV: 92% NPV: 83%
Chang et al., 2008[54]	72	Docetaxel every 3 wk × 4 cycles	Core biopsies; RT-PCR	Complete clinical response: 12/72 (17%) Partial response[a]: 41/72 (57%) Stable disease: 17/72 (24%) Progressive disease[b]: 2/72 (3%)	14-gene signature (derived from above 92-gene signature) significantly related to complete clinical response ($p < 0.05$)
Iwao-Koizumi et al., 2005[68]	Training set: 44 Validation set: 26	Docetaxel every 3 wk × 4 cycles	Incisional biopsy or vacuum-assisted core needle biopsy; ATAC-PCR	Responders (clinical response and partial response): 24/44 Nonresponders (no change and progressive disease): 22/44	85-gene signature: predictive accuracy: 80.7% sensitivity: 91.7% specificity: 71.4% PPV: 73.3% NPV: 90.9%

a Partial response = decrease in unidimensional size by at least 30%.

b Progressive disease = increase in unidimensional size by more than 25%.

ATAC-PCR = adapter-tagged competitive PCR; **NPV** = negative predictive value; **PPV** = positive predictive value; **RT-PCR** = reverse transcriptase PCR.

($p < 0.0001$ and $p = 0.01$, respectively). However, basal-like and HER2+/ER− subtypes had worse distant disease-free survival ($p = 0.04$) and overall survival ($p = 0.02$) than the luminal subtype. The authors believed this was due to a higher relapse rate in non-pCR basal-like and HER2+/ER− subtypes.

Two other studies created a combined chemotherapy regimen sensitivity signature from individual chemotherapy sensitivity signatures. Bonnefoi et al.[63] concluded that FEC (fluorouracil, epirubicin, and cyclophosphamide) and TET (docetaxel, epirubicin, and docetaxel) regimen-specific signatures are significantly predictive of pCR. In addition, Salter et al.[65] generated the combined TFAC (paclitaxel, fluorouracil, adriamycin [doxorubicin], cyclophosphamide) signature from NCI-60 individual chemotherapy sensitivity signatures. The combined TFAC signature predicted response or pCR. The signature's probability of sensitivity in the responders was higher than in the nonresponders ($p = 0.002$).

Predictive signatures of response to taxane alone have also been reported (table IV). Chang et al.[48,67] evaluated the predictive utility of gene signatures in response to docetaxel. They identified a 92-gene profile, which included the 21-gene signature that was predictive of docetaxel sensitivity and resistance. They later derived a 14-gene signature from the 92-gene signature that significantly predicted complete clinical response in patients treated with docetaxel ($p < 0.05$). Iwao-Koizumi et al.[68] also developed an 85-gene signature that is predictive of response to docetaxel.

4. Conclusions

Gene expression profiling has the potential to revolutionize prognostication and treatment selection for individual patients, based on tumor-specific prognostic and predictive signatures. However, before many of these prognostic and predictive signatures are ready for practical clinical application, large numbers of clinically homogeneous patients must prospectively participate in the validation processes, prognostic and predictive accuracy must be compared between tests and when combined with each other, and more must be understood about how to incorporate these tests into the decision-making process of breast cancer management. The development of several prognostic and predictive signatures and the recent commercial availability of three signatures are promising for an era of treatment tailored to patients' specific breast cancer profiles.

Acknowledgments

Lara Dunn is funded through a FOCUS Medical Student Fellowship in Women's Health, supported by a Bertha Dagan Berman Award. The authors have no conflicts of interest that are directly related to the content of this review.

© 2009 Adis Data Information BV. All rights reserved.

References

1. Clark GM, Osborne CK, McGuire WL. Correlations between estrogen receptor, progesterone receptor, and patient characteristics in human breast cancer. J Clin Oncol 1984; 2: 1102-9
2. Ravdin P. Prognostic factors in breast cancer. Chicago (IL): ASCO education book 1997; 217-27
3. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005; 365: 1687-717
4. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987; 235: 177-82
5. Paik S, Hazan R, Fisher ER, et al. Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: prognostic significance of erbB-2 protein overexpression in primary breast cancer. J Clin Oncol 1990; 8: 103-12
6. Van de Vijver MJ, Mooi WJ, Wisman P, et al. Immunohistochemical detection of the neu protein in tissue sections of human breast tumors with amplified neu DNA. Oncogene 1988; 2: 175-8
7. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002; 20: 719-26
8. Buzdar AU, Ibrahim NK, Francis D, et al. Significantly higher pathologic complete remission rate after neoadjuvant therapy with trastuzumab, paclitaxel, and epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2-positive operable breast cancer. J Clin Oncol 2005; 23: 3676-85
9. National Surgical Adjuvant Breast and Bowel Project, US National Cancer Institute. Initial results of the study of tamoxifen and raloxifene (STAR) release: osteoporosis drug raloxifene shown to be as effective as tamoxifen in preventing invasive breast cancer [media release]. 2006 Jun 21 [online]. Available from URL: http://www.cancer.gov/newscenter/pressreleases/starresultsapr172006 [Accessed 2009 Apr 22]
10. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med 2005; 353: 1659-72
11. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med 2005; 353: 1673-84
12. Harris LN, Liotcheva V, Broadwater G, et al. Comparison of methods of measuring HER-2 in metastatic breast cancer patients treated with high-dose chemotherapy. J Clin Oncol 2001; 19: 1698-706
13. Muss HB, Thor AD, Berry DA, et al. c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. N Engl J Med 1994; 330: 1260-6
14. Paik S, Bryant J, Park C, et al. erbB-2 and response to doxorubicin in patients with axillary lymph node-positive, hormone receptor-negative breast cancer. J Natl Cancer Inst 1998; 90: 1361-70
15. Paik S, Bryant J, Tan-Chiu E, et al. HER2 and choice of adjuvant chemotherapy for invasive breast cancer: National Surgical Adjuvant Breast and Bowel Project Protocol B-15. J Natl Cancer Inst 2000; 92: 1991-8
16. Balslev I, Axelsson CK, Zedeler K, et al. The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). Breast Cancer Res Treat 1994; 32 (3): 281-90
17. Singletary SE, Allred C, Ashley P, et al. Staging system for breast cancer: revisions for the 6th edition of the AJCC Cancer Staging Manual. Surg Clin North Am 2003; 83: 803-19
18. Acharya CR, Hsu DS, Anders CK, et al. Gene expression signatures, clinicopathological features, and individualized therapy in breast cancer. JAMA 2008; 299: 1574-87
19. Tavtigian SV, Pierotti MA, Borresen-Dale AL. International Agency for Research on Cancer workshop on 'Expression array analyses in breast cancer taxonomy'. Breast Cancer Res 2006; 8: 303
20. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000; 406: 747-52
21. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98: 10869-74
22. Ma XJ, Hilsenbeck SG, Wang W, et al. The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. J Clin Oncol 2006; 24: 4611-9
23. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003; 100: 8418-23
24. Goetz MP, Suman VJ, Ingle JN, et al. A two-gene expression ratio of homeobox 13 and interleukin-17B receptor for prediction of recurrence and survival in women receiving adjuvant tamoxifen. Clin Cancer Res 2006; 12: 2080-7
25. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006; 295: 2492-502
26. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817-26
27. Habel LA, Shak S, Jacobs MK, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. Breast Cancer Res 2006; 8: R25
28. Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol 2008; 26: 4063-71
29. Albain K, Barlow W, Shak S. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal, node-positive, ER-positive breast cancer (SWOG8814/TBCI0100). San Antonio Breast Cancer Symposium; 2007 Dec 13-16; San Antonio (TX)
30. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415: 530-6
31. Van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; 347: 1999-2009
32. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006; 98: 1183-92
33. Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 2004; 5: 607-16
34. Reid JF, Lusa L, De Cecco L, et al. Limits of predictive models using microarray data for breast cancer clinical treatment outcome. J Natl Cancer Inst 2005; 97: 927-30
35. Jansen MP, Sieuwerts AM, Look MP, et al. HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study. J Clin Oncol 2007; 25: 662-8
36. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005; 365: 671-9
37. Foekens JA, Atkins D, Zhang Y, et al. Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. J Clin Oncol 2006; 24: 1665-71
38. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TransBIG multicenter independent validation series. Clin Cancer Res 2007; 13: 3207-14

© 2009 Adis Data Information BV. All rights reserved.

39. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci U S A 2005; 102: 3738-43

40. Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 2007; 356: 217-26

41. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res 2007; 13: 2329-34

42. Foulkes WD, Stefansson IM, Chappuis PO, et al. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst 2003; 95: 1482-5

43. Grushko TA, Blackwood MA, Schumm PL, et al. Molecular-cytogenetic analysis of HER-2/neu gene in BRCA1-associated breast cancers. Cancer Res 2002; 62: 1481-8

44. Esteban J, Baker J, Cronin M. Tumor gene expression and prognosis in breast cancer: multi-gene RT-PCR assay of paraffin-embedded tissue [abstract no. 3416]. Proc Am Soc Clin Oncol 2003; 22 [online]. Available from URL: http://www.asco.org/ASCOv2/Meetings/Abstracts?&vmview=abst_detail_view&confID=23&abstractID=104527 [Accessed 2009 May 15]

45. Cobleigh M, Bitterman P, Baker J. Tumor gene expression predicts distant disease-free survival (DDFS) in breast cancer patients with 10 or more positive nodes: high throughput RT-PCR assay of paraffin-embedded tumor tissues [abstract no. 3415]. Proc Am Soc Clin Oncol 2003; 22 [online]. Available from URL: http://www.asco.org/ASCOv2/Meetings/Abstracts?&vmview=abst_detail_view&confID=23&abstractID=104115 [Accessed 2009 May 15]

46. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817-26

47. Goldhirsch A, Glick JH, Gelber RD, et al. Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. J Clin Oncol 2001; 19: 3817-27

48. Chang JC, Wooten EC, Tsimelzon A, et al. Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. J Clin Oncol 2005; 23: 1169-77

49. Al-Hajj M, Wicha MS, Benito-Hernandez A, et al. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100: 2003; 3983-8

50. Mina L, Soule SE, Badve S, et al. Predicting response to primary chemotherapy: gene expression profiling of paraffin-embedded core biopsy tissue. Breast Cancer Res Treat 2007; 103: 197-208

51. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 2006; 355: 560-9

52. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006; 24: 3726-34

53. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. J Clin Oncol 2005; 23: 7265-77

54. Chang JC, Makris A, Gutierrez MC, et al. Gene expression patterns in formalin-fixed, paraffin-embedded core biopsies predict docetaxel chemosensitivity in breast cancer patients. Breast Cancer Res Treat 2008; 108: 233-40

55. Jerevall PL, Brommesson S, Strand C, et al. Exploring the two-gene ratio in breast cancer: independent roles for HOXB13 and IL17BR in prediction of clinical outcome. Breast Cancer Res Treat 2008; 107: 225-34

56. Goetz MP, Suman VJ, Couch FJ, et al. Cytochrome P450 2D6 and homeobox 13/interleukin-17B receptor: combining inherited and tumor gene markers for prediction of tamoxifen resistance. Clin Cancer Res 2008; 14: 5864-8

57. Zujewski JA, Kamin L. Trial assessing individualized options for treatment for breast cancer: the TAILORx trial. Future Oncol 2008; 4 (5): 603-10

58. Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol 2004; 22: 2284-93

59. Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005; 11: 5678-85

60. Hannemann J, Oosterkamp HM, Bosch CA, et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol 2005; 23: 3331-42

61. Cleator S, Tsimelzon A, Ashworth A, et al. Gene expression patterns for doxorubicin (Adriamycin) and cyclophosphamide (Cytoxan) (AC) response and resistance. Breast Cancer Res Treat 2006; 95: 229-33

62. Dressman HK, Hans C, Bild A, et al. Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. Clin Cancer Res 2006; 12: 819-26

63. Bonnefoi H, Potti A, Delorenzi M, et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncol 2007; 8: 1071-8

64. Modlich O, Prisack HB, Munnes M, et al. Predictors of primary breast cancers responsiveness to preoperative epirubicin/cyclophosphamide-based chemotherapy: translation of microarray data into clinically useful predictive signatures. J Transl Med 2005; 3: 32

65. Salter KH, Acharya CR, Walters KS, et al. An integrated approach to the prediction of chemotherapeutic response in patients with breast cancer. PLoS ONE 2008; 3 (4): e1908

66. Yang SX, Steinberg SM, Nguyen D, et al. Gene expression profile and angiogenic marker correlates with response to neoadjuvant bevacizumab followed by bevacizumab plus chemotherapy in breast cancer. Clin Cancer Res 2008; 14: 5893-9

67. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003; 362: 362-9

68. Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol 2005; 23: 422-31

Correspondence: Dr *Angela DeMichele*, Rowan Breast Center, 14 Penn Tower, 3400 Spruce Street, Philadelphia, PA 19104, USA.
E-mail: angela.demichele@uphs.upenn.edu

© 2009 Adis Data Information BV. All rights reserved.