Breast Cancer RESEARCH

REVIEW

# Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction

Pierre-Emmanuel Colombo[1], Fernanda Milanezi[1], Britta Weigelt*[2] and Jorge S Reis-Filho*[1]

**Abstract**

Breast cancer comprises a collection of diseases with distinctive clinical, histopathological, and molecular features. Importantly, tumors with similar histological features may display disparate clinical behaviors. Gene expression profiling using microarray technologies has improved our understanding of breast cancer biology and has led to the development of a breast cancer molecular taxonomy and of multigene 'signatures' to predict outcome and response to systemic therapies. The use of these prognostic and predictive signatures in routine clinical decision-making remains controversial. Here, we review the clinical relevance of microarray-based profiling of breast cancer and discuss its impact on patient management.

## Introduction

Breast cancer is a heterogeneous and complex disease that encompasses different entities with distinct biological features and clinical outcomes [1-3]. Adjuvant systemic therapies are employed to eradicate potential micrometastatic disease after surgery for early-stage cancers but their management remains challenging in clinical practice. Treatment decisions still are based largely on clinicopathological criteria, including age, tumor size, histological grade, lymph node metastasis, lymphovascular invasion, and estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status. These parameters have been incorporated into guidelines such as those of St. Gallen [4] and the National Institutes of Health consensus or integrated in internet-based decision tools (like Adjuvant! Online [5,6]) to aid clinicians evaluating the risk of distant recurrence and the need for adjuvant chemotherapy (Figure 1). While this approach has improved survival for the average population, it has progressively widened the indications of adjuvant chemotherapy [7]. Currently, approximately 60% of all patients with early breast cancer receive some form of chemotherapy; although all patients will be exposed to the toxicity of these agents, only a minority will benefit from it [7,8]. Reliable prognostic and predictive markers are needed to guide the selection of the most appropriate adjuvant therapies for individual patients with breast cancer. In fact, a shift from defining the cancer patients who should receive chemotherapy on the basis of their prognostic characteristics to defining the patients who are likely to benefit most from this modality of adjuvant treatment is currently taking place.

In the past decade, the development of gene expression profiling using high-throughput microarray-based methods has allowed the concurrent analysis of the expression level for thousands of genes in a tumor sample. These technologies were hailed as a new dawn in cancer biology and oncology practice; however, after the initial wave of enthusiasm, a wave of (over)skepticism followed [9,10]. Fortunately, with the significant number of studies based on gene expression profiling in the last decade and the availability of datasets for reanalyses and meta-analyses, the field of gene expression profiling has matured.

Microarray-based gene expression profiling studies undoubtedly have contributed to our understanding of the heterogeneity and complexity of breast cancer behavior. It was through a series of seminal studies by the Stanford group [11-13] that the breast cancer research community has come to terms with the idea that breast cancer is by no means a single disease and that distinct

*Correspondence: britta.weigelt@cancer.org.uk; jorge.reis-filho@icr.ac.uk
[1]Molecular Pathology Team, Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, 237 Fulham Road, London, SW3 6JB, UK
[2]Signal Transduction Laboratory, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London, WC2A 3LY, UK

BioMed Central © 2011 BioMed Central Ltd

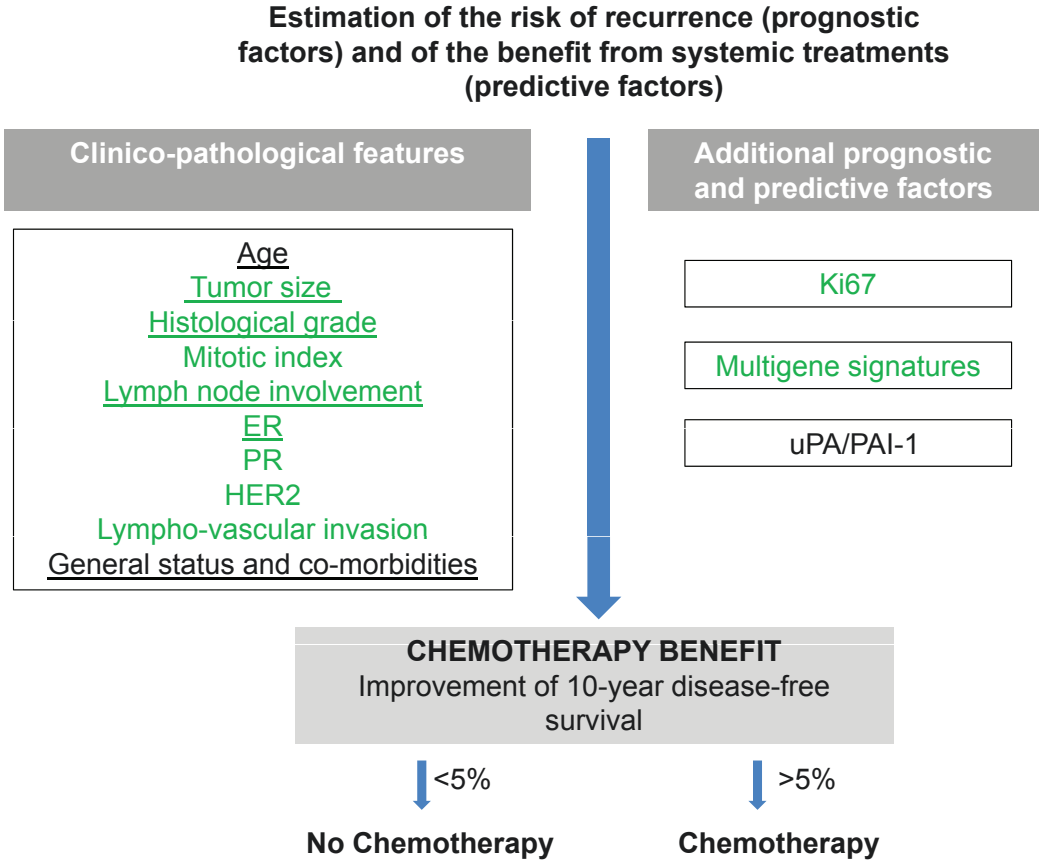


**Figure 1. Clinical decision-making for adjuvant chemotherapy.** Criteria included in the St. Gallen guidelines (green font) and in Adjuvant! Online (underlined) are shown. ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor; uPA/PAI-1, urokinase-type plasminogen activator and plasminogen activator inhibitor-1.

molecular subtypes, often with identical histopathological features, do exist [11]. Moreover, numerous multigene signatures associated with prognosis and response to systemic therapies have emerged [1-3]. Some of these signatures are commercially available (Table 1) and two of them (MammaPrint, Agendia BV, Amsterdam, The Netherlands, and Oncotype DX, Genomic Health, Redwood City, CA, USA) are currently being tested in randomized prospective clinical trials [14,15]. Here, we discuss the potential clinical relevance of gene profiling in breast cancer and its potential impact on patients' clinical care.

## Molecular classification of breast cancer

That breast cancer comprises a heterogeneous and complex group of tumors has been known for decades, and attempts to develop standardized classification systems to account for the diversity of this disease were initiated in the late '60s [16]. Nevertheless, clinical and translational investigators had historically considered breast cancer to be a single group of tumors in the context of clinical trials. The observation that tumors that had similar histopathological characteristics behaved in distinct manners was often used to disregard the histological heterogeneity of breast cancer.

The whole landscape of breast cancer research changed with the publication of seminal, class discovery, microarray-based gene expression profiling studies [11-13], in which the heterogeneity and complexity of breast cancers were rediscovered at the molecular level (Figure 2). To the average 'microarrayer' and bioinformatician, the experiments performed by Perou and colleagues [11] may now sound almost quaint, but in 2000 they had a major impact on how breast cancer was perceived given that they demonstrated that (a) ER-positive and ER-negative breast cancers were fundamentally distinct at the transcriptomic level and (b) breast cancer could be divided into at least five molecular subtypes: luminal A, luminal B, normal breast-like, HER2, and basal-like [12,17] (Figure 2).

Several groups have now demonstrated that ER-positive and ER-negative breast cancers have their prognosis governed by distinct biological processes [18,19]

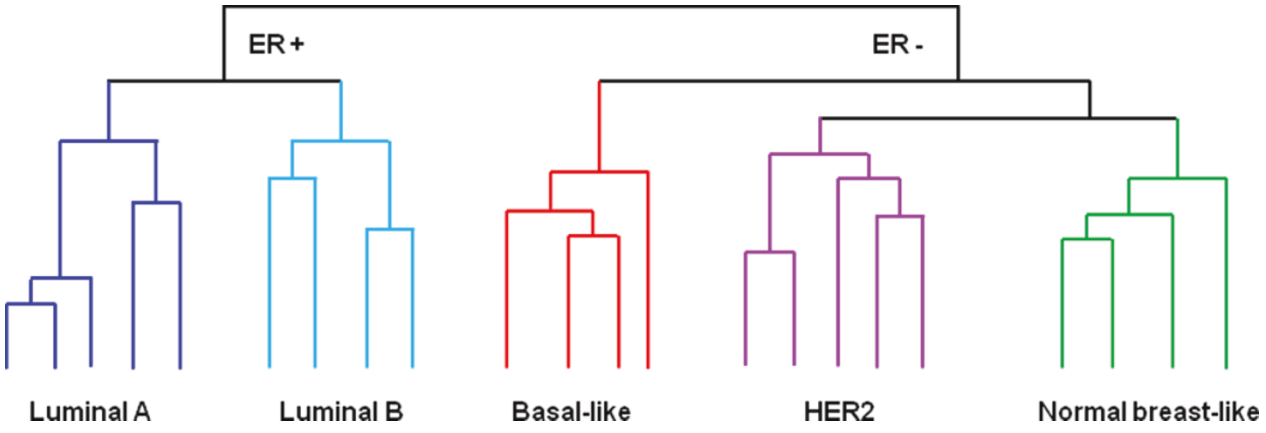


**Figure 2. Schematic illustrations of the five major clusters that represent the molecular subtypes of breast cancer.** Perou and colleagues [11] carried out a cDNA microarray analysis of 38 invasive breast cancers, 1 ductal carcinoma *in situ*, 1 fibroadenoma and 3 normal breast samples, and a number of biological replicates of tumors from the same patients and defined an 'intrinsic gene' list (that is, genes that vary more between tumors from different patients compared with samples from the same tumor/patient). Hierarchical clustering analysis using these 'intrinsic' genes led to the identification of four subtypes: luminal, normal breast-like, human epidermal growth factor receptor 2 (HER2), and basal-like. In subsequent studies, it was demonstrated that similar molecular subtypes of breast cancer could be identified in multiple cohorts and that luminal cancers could be subclassified into two groups (luminal A and B) [12] or three groups (luminal A, B, and C) [13]. The estrogen receptor (ER)-positive branch of the dendrogram contains the luminal tumors, which express low-molecular weight cytokeratins 8/18, ER, and genes associated with an active ER pathway [2,3,11-13,17,26,34]. Luminal A tumors (dark blue) present high levels of expression of ER-activated genes and low proliferation rates and are associated with an excellent prognosis, whereas luminal B breast cancers (light blue) are more often of higher histological grade and have higher proliferation rates and a worse prognosis [2,3,11-13,17,26,34]. The ER-negative branch includes at least three subtypes: normal breast-like, HER2, and basal-like. HER2 tumors (purple) overexpress HER2 and genes associated with the *HER2* amplicon on 17q12 (that is, *GRB7*) and/or the HER2 pathway [2,3,11-13,17,26,34]. Basal-like tumors (red) express genes usually found in normal basal/myoepithelial cells of the breast, including high-molecular weight cytokeratins (5 and 17), caveolins 1 and 2, P-cadherin, nestin, CD44, and EGFR [20]. Morphological and immunohistochemical features of basal-like cancers are similar to those described for tumors arising in *BRCA1* germ-line mutation carriers [20]. The HER2 and basal-like subgroups share an aggressive clinical behavior. Normal breast-like cancers (green) are still poorly characterized [3,22] and there is evidence to suggest that they may constitute an artefact of gene expression profiling associated with a disproportionately high content of normal breast tissue [3,17,26,34].

and that at least some of these subtypes (for example, basal-like) have distinct risk factors, clinical presentation, histological features, response to therapy, and outcome [2,3,20]. These data have led some experts in the field to suggest that traditional clinicopathological features and immunohistochemical markers be replaced by this molecular taxonomy [21].

The initial approach employed for the identification of the molecular subtypes was based on hierarchical clustering analysis. It should be noted, however, that this approach requires large datasets, is to some extent subjective, and cannot be employed for the classification of individual samples prospectively [22-25]. Therefore, 'single sample predictors' (SSPs) were developed on the basis of the correlation between the expression profile of a given sample with the centroids for each molecular subtype (that is, average expression profile of each molecular subtype) [13,17,26]. Over the last decade, three distinct SSPs were developed [13,17,26]. Furthermore, on the basis of this approach, Parker and colleagues [17] developed a quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR)-based or NanoString-based method (PAM50) that can be used to classify formalin-fixed paraffin-embedded (FFPE) samples into the molecular subtypes. Our group [27] and others [28,29] have demonstrated that subtle variations in data normalization and centering, as well as in the proportion of samples from each of the subtypes, may lead to changes in the classification of samples using SSPs. Moreover, independent groups have demonstrated that the classification of tumors into the molecular subtypes, except for the basal-like subtype, is dependent on the SSP used [27,28]. This is best exemplified by the modest agreement in the classification of samples (agreement of 64%, kappa score of 0.527, and 95% confidence interval of 0.456 to 0.597) when a cohort of 295 breast cancers was classified into the molecular subtypes by the authors of the original studies on the molecular classification using SSPs by Sorlie's [13,30] and Perou's [26,31] groups.

Despite the enthusiasm for the use of this molecular taxonomy for the design of clinical trials and routine oncology practice, there are several issues that ought to be considered. First, the subdivision of luminal tumors into A and B is strongly dependent on the SSP used [27] and principally depends on the expression of proliferation-related genes [17,26,32]; there is burgeoning evidence to

demonstrate that the expression of proliferation-related genes in luminal cancers forms a continuum [3,19,33] and that the division of these tumors into two subgroups on the basis of the currently available SSPs [13,17,26] may be artificial. The subclassification of ER-positive breast cancers into subtypes is not only a challenge for the 'intrinsic' subtype classification. In fact, given that proliferation is a continuum in ER-positive cancers and that proliferation is a strong determinant of outcome in this group of tumors, the allocation of ER-positive breast cancer patients into good or poor prognosis by using other microarray-based methods (for example, MammaPrint and genomic grade index) or into low, intermediate, or high histological grade should be considered arbitrary to some extent (see 'Multigene prognostic signatures' section). Second, normal breast-like cancers are now considered by some to be an invalid molecular subtype given that these tumors are likely to constitute an artefact of frozen tissue procurement and representation (that is, samples with a disproportionately high content of normal breast and stromal cells) [3,17,26,27,34,35]. Third, the HER2 (or HER2-enriched) subtype as defined by microarrays does not encompass all cases classified as HER2-positive by clinically validated methods (that is, immunohistochemistry and *in situ* hybridization with methods approved by the US Food and Drug Administration), and not all HER2-positive cancers by clinical methods are classified as HER2 subtype by microarrays [3,17,21,36,37].

The above discrepancies do not invalidate the existence of the 'intrinsic' subtypes. As recently pointed out by Perou and colleagues [38], this is an evolving classification system and PAM50 [17], rather than the SSPs described by Sorlie and colleagues [13] or Hu and colleagues [26], should be employed. With the development of the PAM50 assay, prospective testing of this classification by independent groups will determine its prognostic and predictive power and clinical utility above and beyond the clinicopathological parameters currently available.

The putative histogenetic implications of the molecular subtypes (that is, luminal cancers would originate from luminal cells and basal-like cancers would stem from basal/progenitor cells) [12,13,39-42] have proven incorrect. Although this observation does not have a direct impact on the clinical utility of the 'intrinsic' molecular subtypes, it has led to the assumption that different subtypes of breast cancer would originate from different cell types [13,39-42]. Importantly, there is independent direct evidence to demonstrate that the likeliest cell of origin of basal-like breast cancers lies in the luminal progenitor population rather than the 'basal' population of the normal breast [43,44].

Additional evidence to support the contention that the 'intrinsic' molecular taxonomy remains a working model in development stems from the recent identification of at least three additional molecular subtypes of ER-negative cancers: the 'interferon-rich' subtype [26,45], the 'molecular apocrine' subtype [46-48], and the 'claudin-low' subgroup [35,49] (Figure 2). The 'interferon-rich' subtype, first described by Hu and colleagues [26], is characterized by high expression of interferon-regulated genes, such as *STAT1* [26,45]; the 'molecular apocrine' subtype, which is characterized by activation of androgen receptor signaling, frequently displays *HER2* gene amplification and may be associated with *PTEN* germline mutations [46-48]; and the 'claudin-low' subgroup, which comprises tumors that express low levels or lack of expression of E-cadherin and claudin mRNA, displays an enrichment for the expression of genes often expressed in the process of epithelial-to-mesenchymal transition and immune response genes and allegedly harbors features suggestive of a 'cancer stem cell-like' phenotype [35,49]. Intriguingly, greater than 40% of these cancers do express E-cadherin and claudins at the protein level, despite the low expression levels of these genes by microarray analysis [35]. Importantly, a substantial proportion of tumors classified as of claudin-low subtype by using the cell line-derived SSP described by Prat and colleagues [35] were previously classified as normal breast-like by using other SSPs; these samples may have a disproportionately high content of stromal and normal breast cells. Hence, it remains to be determined whether breast cancers that do express E-cadherin and claudins at the protein level and that were classified as claudin-low by the SSP predictor were not classified as such due to stromal cell contamination. Another point for consideration is the overlap between the transcriptomic characteristics of the claudin-low subtype and those of spindle cell metaplastic breast carcinomas [49,50].

Given the above observations, but despite recent claims that PAM50 models derived from archival formalin-fixed RNA are 'a potential replacement for grade-, hormone receptor-, Ki67-, and HER2-based prognostic models' [21], we argue that the microarray-based gene classification for breast cancer is not yet ready for clinical use in prognostic models or otherwise [1,3,8,27]. In fact, standardization of the definitions and the methodologies for the identification of the molecular subtypes and prospective clinical trials to validate the contribution of the 'intrinsic' subtypes in addition to the current clinicopathological parameters for the management of breast cancer patients are still required [1,3,8,27]. Robust, independently validated methods for the identification of these subtypes are yet to be published.

## Multigene prognostic signatures

### First-generation signatures

The development of microarray-based multigene prognostic classifiers (also known as 'gene signatures') has

been pursued by many groups in the last decade [51-58] with the aim of defining which patients would have such a good prognosis that they could forgo chemotherapy. The first prognostic gene signature [51] consisted of 70 genes and was shown to identify a group of good-prognosis patients with minimal risk of development of distant metastasis within 5 years in patients who were systemic therapy-naïve. In a subsequent study, van de Vijver and colleagues [59] demonstrated that the 70-gene signature was a predictor of outcome independently of the current clinicopathological prognostic markers in a dataset comprising 295 cases (64 cases from the analysis that led to the development of the 70-gene signature and 231 new cases). Importantly, in that [59] and subsequent [60,61] studies, it has been repeatedly demonstrated that the 70-gene signature classifies greater than 95% of ER-negative cancers as poor prognosis and that there is a strong correlation between 70-gene signature-defined poor prognosis and high histological grade. Furthermore, the studies demonstrated that the 70-gene signature would outperform the current methods based on clinicopathological parameters for chemotherapy use [51,59]. This has led to the development of MammaPrint, a commercially available version of the 70-gene signature. Subsequent studies have led to the development of several other prognostic signatures, including the 76-gene signature [54,62] and genomic grade index [55,63-65], which were also shown to be independent predictors of outcome. MammaPrint is currently being tested in the MINDACT (Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid ChemoTherapy) trial [15] (Figure 3), which will determine whether this signature can actually replace clinicopathological parameters for the identification of patients who could be spared from the use of chemotherapy. Table 1 summarizes the prognostic signatures more extensively studied to date. For comprehensive reviews on microarray-based prognostic gene signatures, readers are referred to Sotiriou and Pusztai [2], Weigelt and colleagues [3], and Kim and Paik [66].

In parallel with the development of microarray-based prognostic signatures, Paik and colleagues [52] developed Oncotype DX, a qRT-PCR-based analysis of 21 genes (16 cancer-related and 5 reference genes), which can be used for risk stratification of ER-positive, node-negative breast cancers from patients treated with adjuvant tamoxifen. In contrast to microarray-based predictors, Oncotype DX can be applied to FFPE samples, and this test was developed and validated on the basis of a retrospective analysis of the existing material from two randomized clinical trials (NSABP-B-20 and NSABP-B-14). The signature is based on the expression of genes that are associated with proliferation, ER signaling, HER2, and invasion [52]. The expression of these genes is presented as a recurrence score (RS) that ranges from 0 to 100. These scores provide an estimate of 10-year distant recurrence risk. For clinical use, patients are separated into three categories: low-risk (RS <18), intermediate-risk (RS ≥18 and <31), and high-risk (RS ≥31) [52]. Oncotype DX has been shown to be an independent prognostic factor for patients with ER-positive, node-negative breast cancers treated with tamoxifen and to outperform standard clinicopathological parameters for the prediction of 10-year distant recurrence risk [52]. Oncotype DX has been subsequently evaluated in other populations of breast cancer [67] and shown to be an independent prognostic parameter in patients with ER-positive tumors with up to three positive nodes receiving adjuvant chemotherapy [68] and in postmenopausal patients with ER-positive tumors treated with aromatase inhibitors (that is, anastrozole) [69].

Oncotype DX RSs have also been shown to be correlated with the benefit patients derive from adjuvant chemotherapy in samples from clinical trials [70-72]. In fact, patients with tumors displaying high RSs despite the poor prognosis derive significantly more benefit from chemotherapy than those with low-RS tumors. In addition, patients with low-RS cancers appear to derive negligible benefit from the addition of chemotherapy to tamoxifen [70,71]. Therefore, Oncotype DX has also been considered a predictive marker of benefit from chemotherapy.

Despite the numerous publications on first-generation signatures, level II evidence to support the prognostic role was achieved only for Oncotype DX; for the remaining signatures, only level III evidence has been obtained so far. Given the level of evidence that has been accrued, Oncotype DX has received approval from the American Society of Clinical Oncology [73] and was included in the National Comprehensive Cancer Network guidelines (Breast Cancer version 1.2011 [74]) as an option to evaluate prognosis and as a complement to clinicopathological features to predict response to chemotherapy for patients with ER-positive, node-negative breast cancer. None of the microarray-based prognostic signatures has been endorsed by these professional bodies.

### Are the first-generation prognostic gene signatures ready for use in clinical practice?

Although the different first-generation signatures described above provide relevant information for outcome prediction, they have yet to be incorporated into clinical practice [1,3,8]. The reasons for this apparent failure are multifactorial, and not a single first-generation signature is currently supported by level I evidence for their prognostic power. This information will be available only after the completion of the two randomized trials, MINDACT [15] and TAILORx (Trial Assigning IndividuaLized Options

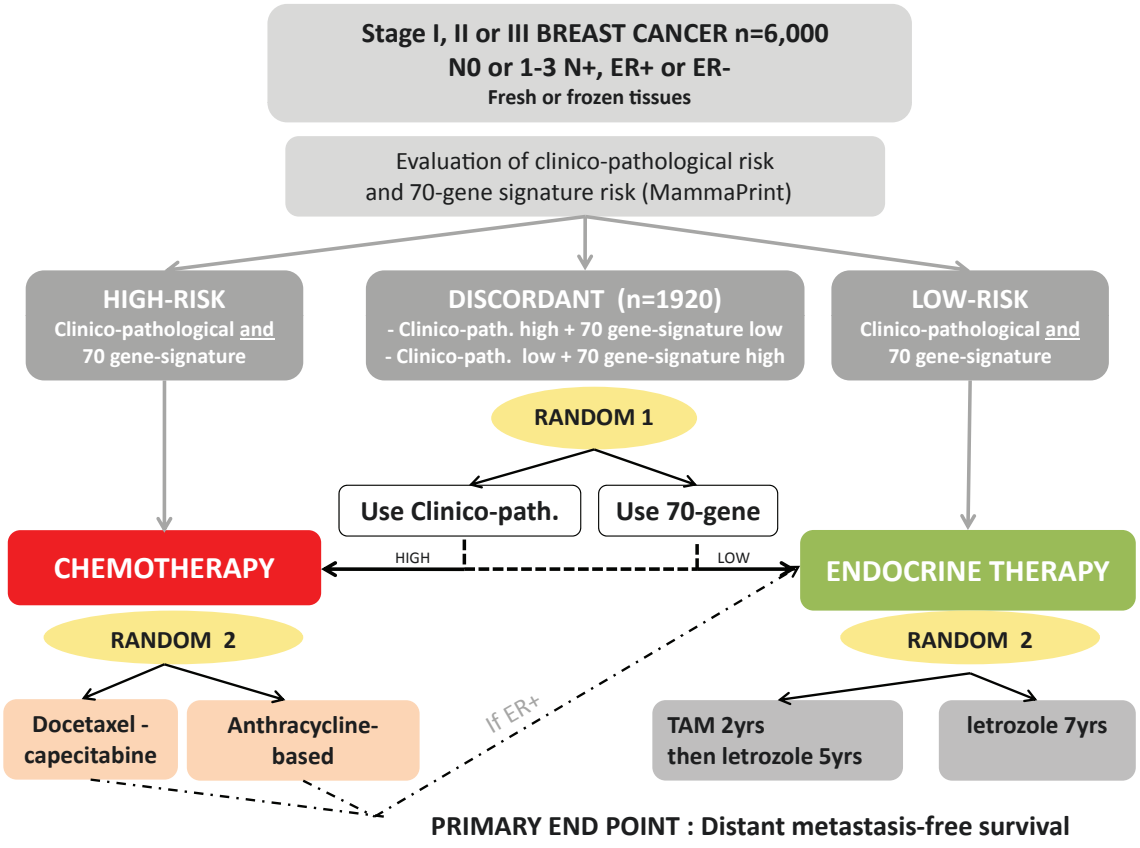


**Figure 3. MINDACT (Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid ChemoTherapy) randomized trial design.** The clinical impact of MammaPrint is being evaluated in MINDACT, a prospective multicenter randomized trial conducted by the European Organization for Research and Treatment of Cancer. The trial compares the recurrence-risk assessment of the 70-gene signature with that provided by Adjuvant! Online in selecting patients for adjuvant chemotherapy. Patients with concordant results are being treated accordingly (high-risk: chemotherapy with or without endocrine therapy, depending on estrogen receptor (ER) status; low-risk: hormonal therapy if ER-positive without chemotherapy). Discordant cases are being randomly assigned to receive adjuvant therapy on the basis of either clinicopathological or 70-gene signature risk assessment. Launched in 2006, the trial intends to confirm the validity of the signature and demonstrate that its clinical use would reduce the number of patients receiving unnecessary treatments, but the results will probably take years to be revealed. Clinico-path, clinicopathological; N, lymph node; N0, lymph node-negative; RANDOM, randomization; TAM, tamoxifen; yrs, years.

for Treatment Rx) [14] (Figures 3 and 4), which evaluate the genomic signatures MammaPrint and Oncotype DX, respectively.

First-generation signatures have been shown not to be stable in terms of the list of genes they are composed of [75,76]; however, comparative studies and meta-analyses have demonstrated that, despite having a negligible overlap in their constituent genes, the first-generation signatures tend to have similar performance and show a relatively good concordance in their prognostic classification, identifying similar but not identical subgroups of patients with poor prognosis [31,33,77].

The ability of these signatures to determine prognosis seems to be directly correlated to the assessment of proliferation-/cell cycle-related genes [18,33]. The fact that these first-generation signatures arguably are mere surrogates of proliferation poses some important problems for their use. First, given that proliferation has been shown to be prognostic in ER-positive disease and not in ER-negative cancers, first-generation signatures are applicable only for the prognostication of patients with ER-positive and HER2-negative breast cancers [18,54,60,61]. As the expression level of proliferation-related genes in ER-positive cancers has been demonstrated to follow a continuum rather than a bimodal distribution, the subdivision of ER-positive cancers into good-prognosis (that is, luminal A) and poor-prognosis (that is, luminal B) groups is artificial [18,33]. In fact, the continuous nature of the Oncotype DX RS is more representative of the ranges of prognosis of patients with ER-positive disease. It should be noted, however, that this approach for clinical decision-making may be problematic. For instance, the prognostication and management of patients with an intermediate RS remain unclear, and

**Table 1. Prognostic multigene signatures in breast cancer commercially available or in commercial development**

| Signature | MammaPrint [51,59] | Oncotype DX [52] | Theros/MGI [53,56-58] | MapQuant DX/ simplified [55] | Veridex 76-gene [54] |
|---|---|---|---|---|---|
| Commercially available/Provider | Yes/Agendia BV (Amsterdam, The Netherlands) | Yes/Genomic Health (Redwood City, CA, USA) | Yes/ bioTheranostics, Inc. (San Diego, CA, USA) | Yes/Ipsogen Inc. (Stamford, CT, USA) | No/Johnson & Johnson (New Brunswick, NJ, USA) |
| Study population | ER+ and ER–, N0, <5 cm diameter, age <55 years | ER+, N0, TAM treated | ER+, N0 | ER+ and ER–, N0 and N+ | ER+ and ER– |
| Assay | 70-gene signature | 21-gene Recurrence Score | 2-gene HOXB13:IL17R/ molecular-grade index | 97-gene signature/8-gene PCR | 76-gene signature |
| Platform | Microarray (Agilent Technologies, Inc., Santa Clara, CA, USA) | RT-PCR | RT-PCR | Microarray (Affymetrix, Santa Clara, CA, USA)/ RT-PCR | Microarray (Affymetrix) |
| Tissue type | Frozen or stabilized mRNA | FFPE | FFPE | Frozen/FFPE | Frozen |
| Prognostic value in other populations | Age 55-70 years, 1-3 N+, N0 and N+, HER2+ | ER+ and 1-3 N+, ER+ postmenopausal receiving aromatase inhibitors | - | ER+ receiving aromatase inhibitors | |
| Predictive value | Neoadjuvant and adjuvant CT (poor signature) | Neoadjuvant and adjuvant CT [71] (high-RS), response to TAM (low-RS) | Resistance to TAM (high-ratio) | Response to neoadjuvant CT (high-risk) | Response to TAM (high-risk patients) |
| Indication | Prognostic in N0, <5 cm diameter, stage I/II BC, age <61 years | Prediction of recurrence risk in ER+ and N0 BC treated with TAM | Prognostic in ER+ BC, prediction of response to TAM | Molecular grading, for ER+, histological grade II BC | Prognostic in ER+ BC |
| Level of evidence | III | II | III | III | III |
| FDA approval | Yes | No | No | No | No |
| Randomized trial | MINDACT | TAILORx | - | - | - |
| Availability | Europe and USA | Europe and USA | USA | Europe | - |

BC, breast cancer; CT, chemotherapy; ER; estrogen receptor status (+ or –); FDA, US Food and Drug Administration; FFPE, formalin-fixed paraffin-embedded; HER2, human epidermal growth factor receptor 2; HOXB13, homeobox 13; IL-17BR, interleukin-17B receptor; MGI, molecular grade index; MINDACT, Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid ChemoTherapy; N+, lymph node-positive; N0, lymph node-negative; PCR, polymerase chain reaction; RS, recurrence score; RT-PCR, reverse transcriptase-polymerase chain reaction; TAILORx, Trial Assigning IndividuaLized Options for Treatment Rx; TAM, tamoxifen.

up to 40% to 60% of clinically intermediate-risk patients (that is, breast cancers combining ER-positive, HER2-negative, and grade II status) are allocated to the intermediate-risk RS group [78]. Therefore, the actual contribution of Oncotype DX to the management of this particular group of patients remains to be elucidated [78]. The lack of prognostic power of first-generation prognostic signatures in ER-negative breast cancer and their association with proliferation in ER-positive breast cancer have brought to the forefront of cancer research the limitations of histological grading. In a way akin to first-generation prognostic gene signatures, histological grade is not prognostic in ER-negative disease and is strongly associated with proliferation [18,79]. It should be noted, however, that the levels of intra- and inter-observer agreement of histological grade remain sub-optimal, despite the numerous efforts to implement a standardized histological grading system [79]. It could be argued, on the basis of the above observations, that the major contribution of first-generation prognostic gene signatures is to provide a standardized proliferation assay for breast cancer.

A second limitation of the first-generation prognostic signatures stems from the fact that most of them were developed to predict short-term distant recurrence (<5 years) and were shown to have a strong 'time dependence' and a reduced prognostic value after 5 to 10 years of follow-up [61,80]. Hence, these signatures may represent merely early distant recurrence surrogates and are unable to predict late relapses with the same accuracy. Thus, there is still a need to develop signatures that could identify patients who have a higher risk of late relapse and who may benefit from prolonged therapy.

Another important consideration in relation to the currently available first-generation prognostic signatures is that they were derived on the basis of the analysis of tissue samples with varying contents of neoplastic cells, stromal cells, inflammatory infiltrate, and normal breast tissue. There is evidence to suggest that the percentage of non-neoplastic cells has a substantial impact on the final expression profile of a tumor and on the ability to derive biologically meaningful prognostic signatures [81]. It should be noted that, although stromal cells and inflammatory infiltrate may be integral parts of the expression profile of a tumor and provide important prognostic and predictive information, most studies employed samples with percentages of stromal cells, inflammatory infiltrate, and normal breast tissue ranging from 0% to 50%.

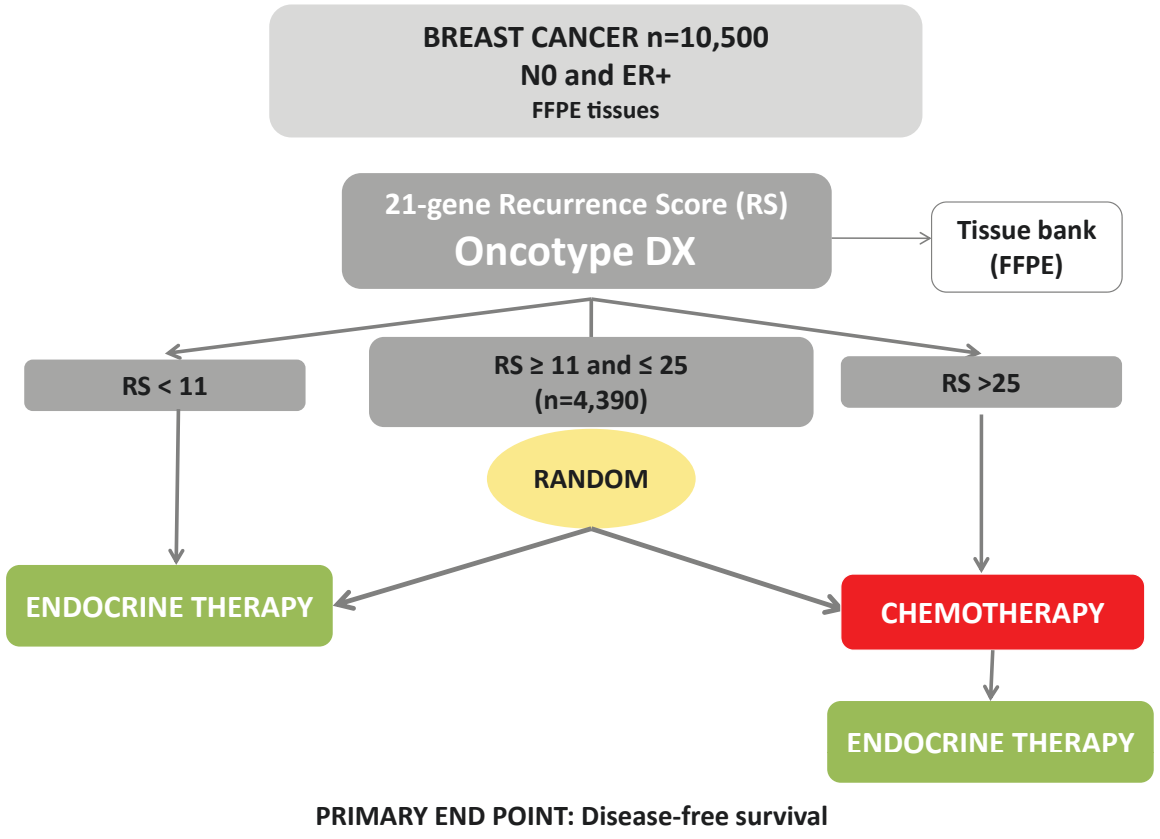


**Figure 4. TAILORx (Trial Assigning IndividuaLized Options for Treatment Rx) randomized trial design.** Oncotype DX is being tested in TAILORx, a prospective randomized phase III trial coordinated by the Breast Cancer Intergroup. The purposes of this trial are to confirm the prognostic value of the 21-gene recurrence score (RS), to determine the optimal management of patients with intermediate-RS, and to refine the utility of the assay in clinical practice. The accrual was recently completed and the first results will be disclosed in 2013. Patients with estrogen receptor (ER)-positive, node-negative breast cancers were eligible and were separated in three study groups according to their RS. High-RS patients (RS >25) received chemotherapy plus endocrine therapy, whereas low-RS patients (RS <11) were assigned to endocrine therapy alone. Patients with intermediate-RS (RS = 11 to 25) were randomly assigned to receive either hormonal therapy alone or hormonal therapy plus chemotherapy. To minimize potential under-treatment in both the high-risk and the randomly assigned groups, the RS ranges for TAILORx were different from those originally defined (11 to 25 instead of 18 to 31). FFPE, formalin-fixed paraffin-embedded; N0, lymph node-negative; RANDOM, randomization.

It remains to be determined whether repeated samples of the same tumor with drastically different percentages of neoplastic cells (for example, 50% versus 100%) would be allocated to the same prognostic subgroup consistently. Therefore, methods to estimate the non-neoplastic cell content of samples or tissue microdissection to standardize the proportion of neoplastic/non-neoplastic cells would be desirable in the development of new microarray-based classifiers and implementation of currently available gene expression signatures.

Despite the initial claims that these signatures would replace current clinicopathological parameters for the management of patients with breast cancer, clinicopathological variables have been shown to add prognostic information independent from that offered by first-generation signatures [1-3]. Therefore, these gene signatures should be perceived as ancillary tools that complement current methods based on the clinicopathological features of the tumors rather than as a replacement for them [1-3]. Importantly, the additional prognostic information provided by first-generation signatures appears to be limited when clinicopathological parameters are analyzed in a centralized fashion with standardized methods (that is, centralized reassessment of histological grade and standardized assessment of ER, PR, HER2, and proliferation rate as defined by Ki67 immunohistochemical analysis) [82]. Therefore, the true contribution of the commercially available first-generation signatures beyond tumor morphology and immunohistochemistry remains to be determined [8].

Recently, 'second-generation' signatures specific for the distinct subtypes of breast cancers have been reported by studying breast cancer microenvironment or host immune response [1,83-87]. Immune response-related signatures

have been shown to be potential prognosticators in ER-negative or triple-negative breast cancers [83-85]. Although these signatures are promising, additional evidence in support of the use of these signatures as potential predictors of outcome is still required.

## Multigene predictive signatures

Beyond prognostic classifiers, an important challenge is to provide physicians with biomarkers that could predict the response or lack of response to treatments and determine the most effective regimen for a specific patient or subgroup of patients. In clinical practice, only ER and HER2 are currently used as predictive markers for the selection of patients likely to respond to endocrine therapy and trastuzumab, respectively.

In addition to Oncotype DX, whose RSs have been shown to be associated with benefit from the addition of chemotherapy to tamoxifen, other prognostic signatures were also shown to have predictive value for the incremental benefit of chemotherapy [1-3,65,88,89]. However, unlike Oncotype DX, the predictive power of MammaPrint [88,89] and genomic grade index [65] have only been tested in retrospective datasets from patients treated with multidrug chemotherapy regimens.

### Gene expression signatures and response to chemotherapy

With the clinical need for predictive markers for specific chemotherapy agents and multidrug regimens, several groups have developed multigene signatures specifically designed to predict response in patients receiving either chemotherapy or endocrine therapy. Using supervised approaches, several studies have attempted to identify multigene signatures of response to chemotherapy by comparing gene expression profiles between high-sensitivity and low-responsiveness tumors [90-93]. The majority of the studies focused on neoadjuvant chemotherapy and, by means of microarrays or RT-PCR, analyzed tumor samples obtained from biopsies taken at diagnosis before initiation of chemotherapy (Table 2). Chemotherapy sensitivity usually was estimated with rate of pathological complete response to neoadjuvant therapy (pCR) as a surrogate of long-term benefit from the treatment. For example, the MD Anderson Cancer Center group developed a 30-gene signature in 82 breast cancer patients receiving T/FAC chemotherapy (paclitaxel, fluorouracil, doxorubicin, cyclophosphamide) [90,92]. This DLDA-30 predictor was then validated in 51 independent patients and predicted pCR probability with higher sensitivity and negative predictive value than clinical variables based on age, grade, and ER status [92]. The accuracy of this predictor was confirmed in an independent study [94]. Despite these interesting preliminary results, the accuracy of the 30-gene predictor was not found in a recent study in which it was not an independent predictor of pCR after multivariate analysis and did not perform better than clinical variables, questioning its potential utility in the clinical setting [95].

An alternative attempt to predict chemosensitivity to specific chemotherapy regimens was developed with the use of *in vitro* models [96]. The combination of *in vitro* signatures associated with drug sensitivity in cell lines was thought to provide composite signatures that could predict response to multidrug regimens and be translated to patients receiving multidrug chemotherapy [96]. These 'regimen-specific' signatures tested in patients who, as participants in the European Organization for Research and Treatment of Cancer (EORTC) BIG00-01 clinical trial, received TET (docetaxel, epirubicin-docetaxel) or FEC (fluorouracil, epirubicin, and cyclophosphamide) chemotherapy resulted in a validation study published in 2007 [97]. Importantly, problems with the methodology of these studies have been identified [98-100] and serious concerns about the validity of the published results were raised [101,102]. Subsequently, after a series of investigations, the findings derived from *in vitro* studies were considered invalid, and this led to the discontinuation of the clinical trials based on these prediction models. Furthermore, several high-profile publications have recently been retracted.

Another method to develop multigene classifiers of chemosensitivity is based on the use of metagenes (that is, groups of coexpressed genes associated with a small number of biological processes). A retrospective microarray analysis of prospectively collected ER-negative breast cancer samples demonstrated that increased stromal gene expression predicted resistance to FEC chemotherapy [103]. This 'stromal' multigene classifier was subsequently validated in two independent cohorts [103]. Further validation of this metagene is awaited.

Despite the promising initial results, the signatures of chemotherapy sensitivity have so far had limited use in clinical practice. Most of them have been developed in small, convenience cohorts and require further external validation. None of the different predictors of chemosensitivity is commercially available, and additional evidence is still required before they can be implemented in clinical practice. For a detailed discussion of the reasons for the limited success of the predictive signatures available to date, readers are referred to a recent review by Borst and Wessels [102]. On the basis of the design employed in most of the studies, the predictive signatures for multidrug regimens are likely to capture the transcriptomic features of sensitivity/resistance to cytotoxic agents in general. These mechanisms may constitute convergent phenotypes [104] (that is, there are multiple genetic/epigenetic aberrations that may lead to resistance to cytoxic agents). The next generation of signatures ought to focus on specific drugs within a given

**Table 2. Multigene predictors of sensitivity to chemotherapy**

| Authors | Number of cases[a] | Regimen | Chemotherapy | Chemosensitivity evaluation | Technology | Method | Signature | NPV | PPV | Accuracy |
|---|---|---|---|---|---|---|---|---|---|---|
| Chang *et al*. [116] | 24 discovery 6 validation | Neoadjuvant | Docetaxel | Clinical response | cDNA microarray | Supervised | 92 genes | 83% | 92% | 88% |
| Ayers *et al*. [90] | 24 discovery 12 validation | Neoadjuvant | T/FAC | pCR | cDNA microarray | Supervised | 74 genes | 73% | 100% (3/3) | 78% |
| Iwao-Koizumi *et al*. [91] | 44 discovery 26 validation | Neoadjuvant | Docetaxel | Clinical response | High-throughput RT-PCR | Supervised | 85 genes | 90.9% | 73.3% | 80.7% |
| Gianni *et al*. [70] | 89 discovery 92 validation | Neoadjuvant | TA | pCR | qRT-PCR/ DNA microarray | Supervised | 86 genes | - | - | - |
| Hess *et al*. [92] | 82 discovery 51 validation | Neoadjuvant | T/FAC | pCR | cDNA microarray | Supervised | 30 genes | 96% | 52% | 76% |
| Thuerigen *et al*. [93] | 52 discovery 48 validation | Neoadjuvant | G-ET | pCR | cDNA microarray | Supervised | 512 genes | 95% | 64% | 88% |
| Farmer *et al*. [103] | 63 | Neoadjuvant | FEC | pCR | cDNA microarray | Metagene approach | Stromal metagene | 81% | 57% | 65% |

[a]Number of cases in discovery and validation sets. FEC, fluorouracil, epirubicin, and cyclophosphamide; G-ET, gemcitabine, epirubicin, and docetaxel; NPV, negative predictive value; pCR, pathological complete response to neoadjuvant chemotherapy; PPV, positive predictive value; qRT-PCR, quantitative reverse transcriptase-polymerase chain reaction; RT-PCR, reverse transcriptase-polymerase chain reaction; TA, taxanes and anthracycline (that is, paclitaxel and doxorubicin); T/FAC, paclitaxel/fluorouracil, doxorubicin, and cyclophosphamide.

subtype of breast cancer, as the predictors of response to chemotherapy in ER-positive and ER-negative breast cancers appear to be fundamentally different [19]. Furthermore, potential mechanisms of resistance to chemotherapy identified by orthogonal methods (for example, RNA interference screens [105], microarray-based comparative genomic hybridization [106,107], proteomic analyses [108], and hypothesis-driven studies [109]) may be used as the basis for the development of multigene predictive signatures. With the availability of multiple microarray datasets from retrospective cohorts and clinical trials in the public domain, novel signatures derived from analyses using orthogonal methods can be tested in a timely fashion.

### Predictive multigene markers of response to endocrine therapy

ER status has an important negative predictive value for evaluating the response to anti-estrogen therapy. Nevertheless, ER expression alone is not sufficient to predict which ER-positive tumor will respond or be resistant to different modalities of endocrine therapies. Microarray-based gene expression signatures to predict outcome of tamoxifen-treated patients have been developed (Table 3). For example, a 44-gene signature, identified by Jansen and colleagues [110], compared gene expression profiles in patients with advanced ER-positive breast cancers treated by tamoxifen. Other hormone sensitivity tests studying estradiol-induced genes in MCF-7 cell line culture [111] or clusters of correlated genes [112] have also been reported.

More recently, the sensitivity to endocrine therapy (SET) index was developed in a large series of ER-positive breast cancers [113]. The SET index is based on the principle that expression of genes correlated with ER may better predict response to endocrine treatment than ER expression alone [113]. Microarray analysis of a discovery set of ER-positive tumors led to the identification of 165 genes coexpressed with ER; the SET index was devised and applied to a validation cohort to define three categories of sensitivity (low, intermediate, and high). Association between SET and outcome was then analyzed in three types of ER-positive cohorts receiving either adjuvant tamoxifen for 5 years or neoadjuvant chemotherapy followed by endocrine therapy (tamoxifen or aromatase inhibition) or no adjuvant systemic treatment. The SET index was significantly associated with the outcome of patients receiving any type of endocrine treatment (tamoxifen or chemoendocrine treatment) but had no prognostic value in untreated patients. Unlike other multigene signatures evaluating proliferation in ER-positive tumors, the SET index seems to be predictive of benefit from endocrine therapy independently of the inherent prognosis of the tumor. Interestingly, for a potential clinical application, the SET index could identify a subset of tumors associated with an excellent prognosis and no relapse in the tamoxifen-treated group (node-negative and high-SET index tumors) and in the chemoendocrine group (high- and intermediate-SET index). Studies evaluating the clinical relevance of the SET index are warranted to expand its indications in clinical practice.

**Table 3. Multigene predictors of response to endocrine treatment**

| Authors | Signature | Number of cases[a] | Treatment | Context | Method | Platform |
|---|---|---|---|---|---|---|
| Jansen *et al.* [110] | 44 genes | 48 training set,<br>66 validation set | TAM | Recurrent breast cancer | Top-down (response vs. progression) | Microarray (local) |
| Loi *et al.* [112] | Tamoxifen predictor (181 genes) | 99 training set,<br>69 validation set<br>87 validation set | TAM | Adjuvant treatment | Top-down (relapse vs. no relapse) | Microarray (Affymetrix, Santa Clara, CA, USA) |
| Oh *et al.* [111] | Estrogen-regulated genes (822 genes) | Cell lines + 65 discovery set,<br>60 validation set (1)<br>90 validation set (2)<br>250 validation set (3) | Heterogeneous<br>TAM (1)<br>TAM +/– NAC (2)<br>Heterogeneous (3) | Adjuvant treatment | Bottom-up (estrogen induced gene expression changes in MCF-7 cell lines) | Microarray (Agilent Technologies, Inc., Santa Clara, CA, USA) |
| Symmans *et al.* [113] | SET index (165 genes) | 437 training set,<br>225 + 298 validation set (1)<br>122 validation set (2)<br>208 + 133 validation set (3) | Heterogeneous<br>TAM (1)<br>NAC + TAM or AI (2)<br>Non-treated (3) | Adjuvant treatment | Bottom-up (genes coexpressed with estrogen receptor) | Microarray (Affymetrix) |

[a]Number of cases in training and validation sets. AI, aromatase inhibitor; NAC, neoadjuvant chemotherapy; SET, sensitivity to endocrine therapy; TAM, tamoxifen.

### Predictors for specific targeted therapies

To date, only a few gene signatures have been developed to predict the response to specific targeted therapies in breast cancer. Recently, Loi and colleagues [114] reported promising results focusing on *PIK3CA* (phosphoinositide-3-kinase, catalytic) gene mutations and the PI3K-AKT-mTOR signaling pathway targeted by PI3K/mTOR (mammalian target of rapamycin) inhibitors. By analysis of gene expression from 1,800 breast cancers, a gene expression signature associated with *PIK3CA* mutation was developed (PIK3CA-GS). The signature predicted *PIK3CA* mutations in two independent datasets of breast cancers and was shown to identify good-prognosis patients in the ER-positive, HER2-negative breast cancer subgroup even in the case of highly proliferative tumors. The PIK3CA-GS was negatively correlated with mTORC1 signaling, making it a potential predictor of response to PI3K/mTOR inhibitors like rapamycin, rapamycin analogs, or dual kinase inhibitors. Breast cancer cell lines with high PIK3CA-GS were confirmed to be resistant to rapamycin [114]. This approach exemplifies the potential use of microarrays as potential predictive markers for tailored therapies.

## Conclusions

Microarray-based gene expression profiling analysis has undoubtedly had a dramatic impact on our understanding of breast cancer biology by bringing the concept of the heterogeneity of breast cancer to the forefront of breast cancer research and clinical practice. In fact, it is currently inconceivable to consider ER-positive and ER-negative breast cancers to be a single disease. However, how the information derived from the classification of breast cancer into the current molecular subtypes [17] will be used for breast cancer patient management remains unclear. First-generation prognostic signatures have led to the realization of the importance of proliferation for the prognostication of patients with ER-positive cancers [1-3]. However, despite the resources allocated to their development and validation, prognostic signatures have proven to add limited information to prognostic models based on clinicopathological parameters and standardized assessment of ER, PR, HER2, and proliferation. Gene signatures predictive of response to specific chemotherapy regimens have proven elusive. With the development of massively parallel sequencing technologies, it has become possible to determine the repertoire of genetic aberrations a tumor harbors in a single experiment. Given the successful use of genetic information as predictive markers for the use of targeted therapies in breast cancer (for example, *HER2* amplification as a predictive marker for anti-HER2 agents) and tumors from other sites (for example, *KIT* and *PDGFRA* [platelet-derived growth factor receptor alpha] mutations as predictive markers of response to imatinib mesylate in gastrointestinal stromal tumors; *EML4-ALK* gene rearrangements as predictive markers of ALK inhibitors in non-small cell lung cancer), it is plausible that the next generation of classifiers based on sequencing information may have a greater impact on our ability to successfully stratify breast cancer patients into predictive subgroups [115]. Integrative approaches combining genetic, transcriptomic, and proteomic information are likely to lead to breast cancer classification systems that better reflect the biology of the disease, and are more clinically relevant [1]. Although the deluge of high-throughput data will most certainly be a formidable challenge for the breast cancer research community, our ability to characterize tumors at an unprecedented level of detail will undoubtedly lead to novel paradigms for stratified medicine and tailored therapies.

**Abbreviations**
ALK, anaplastic lymphoma kinase; ER, estrogen receptor; FEC, fluorouracil, epirubicin, and cyclophosphamide; FFPE, formalin-fixed paraffin-embedded; HER2, human epidermal growth factor receptor 2; MINDACT, Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid ChemoTherapy; mTOR, mammalian target of rapamycin; NSABP, National Surgical Adjuvant Breast and Bowel Project; pCR, pathological complete response to neoadjuvant therapy; PIK3CA, phosphoinositide-3-kinase (catalytic); PR, progesterone receptor; qRT-PCR, quantitative reverse transcriptase-polymerase chain reaction; RS, recurrence score; RT-PCR, reverse transcriptase-polymerase chain reaction; SET, sensitivity to endocrine therapy; SSP, single sample predictor.

**Competing interests**
The authors declare that they have no competing interests.

**Acknowledgments**
JSR-F and P-EC are funded in part by Breakthrough Breast Cancer. BW is funded by a Cancer Research UK postdoctoral fellowship. P-EC is funded by the Val d'Aurelle Anticancer Centre (Montpellier, France). The authors are grateful to Paul Wilkerson and Violetta Barbashina for the critical reading of the manuscript.

Published: 27 June 2011

**References**
1. Reis-Filho JS, Weigelt B, Fumagalli D, Sotiriou C: **Molecular profiling: moving away from tumor philately.** *Sci Transl Med* 2010, **2:**47ps43.
2. Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *N Engl J Med* 2009, **360:**790-800.
3. Weigelt B, Baehner FL, Reis-Filho JS: **The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade.** *J Pathol* 2010, **220:**263-280.
4. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ: **Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009.** *Ann Oncol* 2009, **20:**1319-1329.
5. **Adjuvant! Online homepage** [https://www.adjuvantonline.com].
6. Mook S, Schmidt MK, Rutgers EJ, van de Velde AO, Visser O, Rutgers SM, Armstrong N, van't Veer LJ, Ravdin PM: **Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study.** *Lancet Oncol* 2009, **10:**1070-1076.
7. Schmidt M, Victor A, Bratzel D, Boehm D, Cotarelo C, Lebrecht A, Siggelkow W, Hengstler JG, Elsasser A, Gehrmann M, Lehr HA, Koelbl H, von Minckwitz G, Harbeck N, Thomssen C: **Long-term outcome prediction by clinicopathological risk classification algorithms in node-negative breast cancer--comparison between Adjuvant!, St Gallen, and a novel risk algorithm used in the prospective randomized Node-Negative-Breast Cancer-3 (NNBC-3) trial.** *Ann Oncol* 2009, **20:**258-264.
8. Weigelt B, Reis-Filho JS: **Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry.** *Breast Cancer Res* 2010, **12 Suppl 4:**S5.
9. Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, Culhane AC, Falchi M, Furlanello C, Game L, Jurman G, Mangion J, Mehta T, Nitzberg M, Page GP, Petretto E, van Noort V: **Repeatability of published microarray gene expression analyses.** *Nat Genet* 2009, **41:**149-155.
10. Ransohoff DF: **Rules of evidence for cancer molecular-marker discovery and validation.** *Nat Rev Cancer* 2004, **4:**309-314.
11. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406:**747-752.
12. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98:**10869-10874.
13. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100:**8418-8423.
14. Sparano JA, Paik S: **Development of the 21-gene assay and its application in clinical practice and clinical trials.** *J Clin Oncol* 2008, **26:**721-728.
15. Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ: **Clinical application of the 70-gene profile: the MINDACT trial.** *J Clin Oncol* 2008, **26:**729-735.
16. Weigelt B, Reis-Filho JS: **Histological and molecular types of breast cancer: is there a unifying taxonomy?** *Nat Rev Clin Oncol* 2009, **6:**718-730.
17. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27:**1160-1167.
18. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, Delorenzi M, Piccart M, Sotiriou C: **Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes.** *Clin Cancer Res* 2008, **14:**5158-5165.
19. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Ya-Hui Shiang C, Santarpia L, Matsuoka J, Hortobagyi GN, Symmans WF, Holmes FA, O'Shaughnessy J, Hellerstedt B, Pippen J, Andre F, Simon R, Pusztai L: **Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer.** *J Natl Cancer Inst* 2011, **103:**264-272.
20. Foulkes WD, Smith IE, Reis-Filho JS: **Triple-negative breast cancer.** *N Engl J Med* 2010, **363:**1938-1948.
21. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MC, Mardis ER, Perou CM, Bernard PS, Ellis MJ: **A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer.** *Clin Cancer Res* 2010, **16:**5222-5232.
22. Pusztai L, Mazouni C, Anderson K, Wu Y, Symmans WF: **Molecular classification of breast cancer: limitations and potential.** *Oncologist* 2006, **11:**868-877.
23. Kapp AV, Tibshirani R: **Are clusters found in one dataset present in another dataset?** *Biostatistics* 2007, **8:**9-31.
24. Loi S, Sotiriou C, Buyse M, Rutgers E, Van't Veer L, Piccart M, Cardoso F: **Molecular forecasting of breast cancer: time to move forward with clinical testing.** *J Clin Oncol* 2006, **24:**721-722; author reply 722-723.
25. Mackay A, Weigelt B, Grigoriadis A, Kreike B, Natrajan R, A'Hern R, Tan DS, Dowsett M, Ashworth A, Reis-Filho JS: **Microarray-based class discovery for molecular classification of breast cancer: analysis of interobserver agreement.** *J Natl Cancer Inst* 2011, **103:**662-673.
26. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, Perou CM: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC Genomics* 2006, **7:**96.
27. Weigelt B, Mackay A, A'Hern R, Natrajan R, Tan DS, Dowsett M, Ashworth A, Reis-Filho JS: **Breast cancer molecular profiling with single sample predictors: a retrospective analysis.** *Lancet Oncol* 2010, **11:**339-349.
28. Haibe-Kains B, Culhane A, Desmedt C, Bontempi G, Quackenbush J, Sotiriou C: **Robusteness of breast cancer molecular subtypes identification. Abstract Book of the IMPAKT 2010 Breast Cancer Conference, Brussels, Belgium, 6-8 May 2010.** *Ann Oncol* 2010, **21:**iv49-iv59.
29. Lusa L, McShane LM, Reid JF, De Cecco L, Ambrogi F, Biganzoli E, Gariboldi M, Pierotti MA: **Challenges in projecting clustering results across gene expression-profiling datasets.** *J Natl Cancer Inst* 2007, **99:**1715-1723.
30. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, Dai H, He YD, van't Veer LJ, Bartelink H, van de Rijn M, Brown PO, van de Vijver MJ: **Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival.** *Proc Natl Acad Sci USA* 2005, **102:**3738-3743.
31. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ, Perou CM: **Concordance among gene-expression-based predictors for breast cancer.** *N Engl J Med* 2006, **355:**560-569.
32. Weigelt B, Geyer FC, Reis-Filho JS: **Histological types of breast cancer: how special are they?** *Mol Oncol* 2010, **4:**192-208.
33. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schutz F, Goldstein DR, Piccart M, Delorenzi M: **Meta-analysis of gene expression profiles in breast cancer:**

toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res* 2008, **10:**R65.
34. Peppercorn J, Perou CM, Carey LA: **Molecular subtypes in breast cancer evaluation and management: divide and conquer.** *Cancer Invest* 2008, **26:**1-10.
35. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast Cancer Res* 2010, **12:**R68.
36. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, Hess KR, Stec J, Ayers M, Wagner P, Morandi P, Fan C, Rabiul I, Ross JS, Hortobagyi GN, Pusztai L: **Breast cancer molecular subtypes respond differently to preoperative chemotherapy.** *Clin Cancer Res* 2005, **11:**5678-5685.
37. de Ronde JJ, Hannemann J, Halfwerk H, Mulder L, Straver ME, Vrancken Peeters MJ, Wesseling J, van de Vijver M, Wessels LF, Rodenhuis S: **Concordance of clinical and molecular breast cancer subtyping in the context of preoperative chemotherapy response.** *Breast Cancer Res Treat* 2010, **119:**119-126.
38. Perou CM, Parker JS, Prat A, Ellis MJ, Bernard PS: **Clinical implementation of the intrinsic subtypes of breast cancer.** *Lancet Oncol* 2010, **11:**718-719; author reply 720-711.
39. Prat A, Perou CM: **Mammary development meets cancer genomics.** *Nat Med* 2009, **15:**842-844.
40. Sorlie T: **Introducing molecular subtyping of breast cancer into the clinic?** *J Clin Oncol* 2009, **27:**1153-1154.
41. Stingl J, Caldas C: **Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis.** *Nat Rev Cancer* 2007, **7:**791-799.
42. Sims AH, Howell A, Howell SJ, Clarke RB: **Origins of breast cancer subtypes and therapeutic implications.** *Nat Clin Pract Oncol* 2007, **4:**516-525.
43. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, Asselin-Labat ML, Gyorki DE, Ward T, Partanen A, Feleppa F, Huschtscha LI, Thorne HJ, Fox SB, Yan M, French JD, Brown MA, Smyth GK, Visvader JE, Lindeman GJ: **Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers.** *Nat Med* 2009, **15:**907-913.
44. Molyneux G, Geyer FC, Magnay FA, McCarthy A, Kendrick H, Natrajan R, Mackay A, Grigoriadis A, Tutt A, Ashworth A, Reis-Filho JS, Smalley MJ: **BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells.** *Cell Stem Cell* 2010, **7:**403-417.
45. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C: **An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer.** *Genome Biol* 2007, **8:**R157.
46. Farmer P, Bonnefoi H, Becette V, Tubiana-Hulin M, Fumoleau P, Larsimont D, Macgrogan G, Bergh J, Cameron D, Goldstein D, Duss S, Nicoulaz AL, Brisken C, Fiche M, Delorenzi M, Iggo R: **Identification of molecular apocrine breast tumours by microarray analysis.** *Oncogene* 2005, **24:**4660-4671.
47. Doane AS, Danso M, Lal P, Donaton M, Zhang L, Hudis C, Gerald WL: **An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen.** *Oncogene* 2006, **25:**3994-4008.
48. Banneau G, Guedj M, MacGrogan G, de Mascarel I, Velasco V, Schiappa R, Bonadona V, David A, Dugast C, Gilbert-Dussardier B, Ingster O, Vabres P, Caux F, de Reynies A, Iggo R, Sevenet N, Bonnet F, Longy M: **Molecular apocrine differentiation is a common feature of breast cancer in patients with germline PTEN mutations.** *Breast Cancer Res* 2010, **12:**R63.
49. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, Fridlyand J, Sahin A, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Carey M, Lluch A, Monteagudo C, He X, Weigman V, Fan C, Palazzo J, Hortobagyi GN, Nolden LK, Wang NJ, Valero V, Gray JW, Perou CM, Mills GB: **Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics.** *Cancer Res* 2009, **69:**4116-4124.
50. Weigelt B, Kreike B, Reis-Filho JS: **Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis.** *Breast Cancer Res Treat* 2009, **117:**273-280.
51. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415:**530-536.
52. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *N Engl J Med* 2004, **351:**2817-2826.
53. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, Muir B, Mohapatra G, Salunga R, Tuggle JT, Tran Y, Tran D, Tassin A, Amon P, Wang W, Enright E, Stecker K, Estepa-Sabal E, Smith B, Younger J, Balis U, Michaelson J, Bhan A, Habin K, Baer TM, Brugge J, Haber DA, Erlander MG, Sgroi DC: **A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen.** *Cancer Cell* 2004, **5:**607-616.
54. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA: **Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.** *Lancet* 2005, **365:**671-679.
55. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M, Delorenzi M: **Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.** *J Natl Cancer Inst* 2006, **98:**262-272.
56. Ma XJ, Hilsenbeck SG, Wang W, Ding L, Sgroi DC, Bender RA, Osborne CK, Allred DC, Erlander MG: **The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer.** *J Clin Oncol* 2006, **24:**4611-4619.
57. Jansen MP, Sieuwerts AM, Look MP, Ritstier K, Meijer-van Gelder ME, van Staveren IL, Klijn JG, Foekens JA, Berns EM: **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:**662-668.
58. Ma XJ, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, Harris A, Goss P, Sotiriou C, Erlander M, Sgroi D: **A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer.** *Clin Cancer Res* 2008, **14:**2601-2608.
59. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R: **A gene-expression signature as a predictor of survival in breast cancer.** *N Engl J Med* 2002, **347:**1999-2009.
60. Bueno-de-Mesquita JM, van Harten WH, Retel VP, van't Veer LJ, van Dam FS, Karsenberg K, Douma KF, van Tinteren H, Peterse JL, Wesseling J, Wu TS, Atsma D, Rutgers EJ, Brink G, Floore AN, Glas AM, Roumen RM, Bellot FE, van Krimpen C, Rodenhuis S, van de Vijver MJ, Linn SC: **Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER).** *Lancet Oncol* 2007, **8:**1079-1087.
61. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F, Piccart MJ: **Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer.** *J Natl Cancer Inst* 2006, **98:**1183-1192.
62. Foekens JA, Atkins D, Zhang Y, Sweep FC, Harbeck N, Paradiso A, Cufer T, Sieuwerts AM, Talantov D, Span PN, Tjan-Heijnen VC, Zito AF, Specht K, Hoefler H, Golouh R, Schittulli F, Schmitt M, Beex LV, Klijn JG, Wang Y: **Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer.** *J Clin Oncol* 2006, **24:**1665-1671.
63. Desmedt C, Giobbie-Hurder A, Neven P, Paridaens R, Christiaens MR, Smeets A, Lallemand F, Haibe-Kains B, Viale G, Gelber RD, Piccart M, Sotiriou C: **The Gene expression Grade Index: a potential predictor of relapse for endocrine-treated breast cancer patients in the BIG 1-98 trial.** *BMC Med Genomics* 2009, **2:**40.
64. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, Buyse M, Bontempi G, Delorenzi M, Piccart MJ, Sotiriou C: **Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade.** *J Clin Oncol* 2007, **25:**1239-1246.
65. Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, Kuerer H, Hortobagyi GN, Piccart-Gebhart M, Sotiriou C, Pusztai L: **Genomic grade index is associated with response to chemotherapy in patients with breast cancer.** *J Clin Oncol* 2009, **27:**3185-3191.
66. Kim C, Paik S: **Gene-expression-based prognostic assays for breast cancer.** *Nat Rev Clin Oncol* 2010, **7:**340-347.
67. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J, Blick NT, Greenberg D, Fehrenbacher L, Langholz B,

Quesenberry CP: **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.
68. Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley S, Shak S, Baehner FL, Ravdin PM, Davidson NE, Sledge GW Jr., Perez EA, Shulman LN, Martino S, Sparano JA: **Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features.** *J Clin Oncol* 2008, **26:**4063-4071.
69. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, Quinn E, Dunbier A, Baum M, Buzdar A, Howell A, Bugarini R, Baehner FL, Shak S: **Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study.** *J Clin Oncol* 2010, **28:**1829-1834.
70. Gianni L, Zambetti M, Clark K, Baker J, Cronin M, Wu J, Mariani G, Rodriguez J, Carcangiu M, Watson D, Valagussa P, Rouzier R, Symmans WF, Ross JS, Hortobagyi GN, Pusztai L, Shak S: **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:**7265-7277.
71. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE Jr., Wickerham DL, Wolmark N: **Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer.** *J Clin Oncol* 2006, **24:**3726-3734.
72. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF: **Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial.** *Lancet Oncol* 2010, **11:**55-65.
73. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, Somerfield MR, Hayes DF, Bast RC: **American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer.** *J Clin Oncol* 2007, **25:**5287-5312.
74. **National Comprehensive Cancer Network homepage** [http://www.nccn.org].
75. Michiels S, Koscielny S, Hill C: **Prediction of cancer outcome with microarrays: a multiple random validation strategy.** *Lancet* 2005, **365:**488-492.
76. Ein-Dor L, Kela I, Getz G, Givol D, Domany E: **Outcome signature genes in breast cancer: is there a unique set?** *Bioinformatics* 2005, **21:**171-178.
77. Haibe-Kains B, Desmedt C, Piette F, Buyse M, Cardoso F, Van't Veer L, Piccart M, Bontempi G, Sotiriou C: **Comparison of prognostic gene expression signatures for breast cancer.** *BMC Genomics* 2008, **9:**394.
78. Kelly CM, Krishnamurthy S, Bianchini G, Litton JK, Gonzalez-Angulo AM, Hortobagyi GN, Pusztai L: **Utility of oncotype DX risk estimates in clinically intermediate risk hormone receptor-positive, HER2-normal, grade II, lymph node-negative breast cancers.** *Cancer* 2010, **116:**5161-5167.
79. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR, Palacios J, Richardson AL, Schnitt SJ, Schmitt FC, Tan PH, Tse GM, Badve S, Ellis IO: **Breast cancer prognostic classification in the molecular era: the role of histological grade.** *Breast Cancer Res* 2010, **12:**207.
80. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, Viale G, Delorenzi M, Zhang Y, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris AL, Klijn JG, Foekens JA, Cardoso F, Piccart MJ, Buyse M, Sotiriou C: **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:**3207-3214.
81. Cleator SJ, Powles TJ, Dexter T, Fulford L, Mackay A, Smith IE, Valgeirsson H, Ashworth A, Dowsett M: **The effect of the stromal component of breast tumours on prediction of clinical outcome using gene expression microarray analysis.** *Breast Cancer Res* 2006, **8:**R32.
82. Cuzick J, Dowsett M, Wale C, Salter J, Quinn E, Zabaglo L, Howell A, Buzdar A, Forbes JF: **Prognostic value of a combined ER, PgR, Ki67, HER2 immunohistochemical (IHC4) score and comparison with the GHI recurrence score - results from TransATAC.** *Cancer Res* 2009, **69:**503S-503S (Abstract).
83. Teschendorff AE, Caldas C: **A robust classifier of high predictive value to identify good prognosis patients in ER-negative breast cancer.** *Breast Cancer Res* 2008, **10:**R73.
84. Yau C, Esserman L, Moore DH, Waldman F, Sninsky J, Benz CC: **A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer.** *Breast Cancer Res* 2010, **12:**R85.
85. Rody A, Karn T, Liedtke C, Pusztai L, Ruckhäberle E, Hanker L, Gaetje R, Solbach C, Ahr A, Metzler D, Schmidt M, Müller V, Holtrich U, Kaufmann M: **Identification of a clinically relevant gene signature in triple negative and basal-like breast cancer.** Paper presented at: 33rd Annual San Antonio Breast Cancer Symposium; San Antonio, TX, USA; 11 December 2010. Abstract S5-5.
86. Bergamaschi A, Tagliabue E, Sorlie T, Naume B, Triulzi T, Orlandi R, Russnes HG, Nesland JM, Tammi R, Auvinen P, Kosma VM, Menard S, Borresen-Dale AL: **Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome.** *J Pathol* 2008, **214:**357-367.
87. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M: **Stromal gene expression predicts clinical outcome in breast cancer.** *Nat Med* 2008, **14:**518-527.
88. Knauer M, Mook S, Rutgers EJ, Bender RA, Hauptmann M, van de Vijver MJ, Koornstra RH, Bueno-de-Mesquita JM, Linn SC, van 't Veer LJ: **The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer.** *Breast Cancer Res Treat* 2010, **120:**655-661.
89. Straver ME, Glas AM, Hannemann J, Wesseling J, van de Vijver MJ, Rutgers EJ, Vrancken Peeters MJ, van Tinteren H, Van't Veer LJ, Rodenhuis S: **The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer.** *Breast Cancer Res Treat* 2010, **119:**551-558.
90. Ayers M, Symmans WF, Stec J, Damokosh AI, Clark E, Hess K, Lecocke M, Metivier J, Booser D, Ibrahim N, Valero V, Royce M, Arun B, Whitman G, Ross J, Sneige N, Hortobagyi GN, Pusztai L: **Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer.** *J Clin Oncol* 2004, **22:**2284-2293.
91. Iwao-Koizumi K, Matoba R, Ueno N, Kim SJ, Ando A, Miyoshi Y, Maeda E, Noguchi S, Kato K: **Prediction of docetaxel response in human breast cancer by gene expression profiling.** *J Clin Oncol* 2005, **23:**422-431.
92. Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, Booser D, Theriault RL, Buzdar AU, Dempsey PJ, Rouzier R, Sneige N, Ross JS, Vidaurre T, Gomez HL, Hortobagyi GN, Pusztai L: **Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer.** *J Clin Oncol* 2006, **24:**4236-4244.
93. Thuerigen O, Schneeweiss A, Toedt G, Warnat P, Hahn M, Kramer H, Brors B, Rudlowski C, Benner A, Schuetz F, Tews B, Eils R, Sinn HP, Sohn C, Lichter P: **Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer.** *J Clin Oncol* 2006, **24:**1839-1845.
94. Peintinger F, Anderson K, Mazouni C, Kuerer HM, Hatzis C, Lin F, Hortobagyi GN, Symmans WF, Pusztai L: **Thirty-gene pharmacogenomic test correlates with residual cancer burden after preoperative chemotherapy for breast cancer.** *Clin Cancer Res* 2007, **13:**4078-4082.
95. Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M, Qi Y, Barajas-Figueroa LJ, Souchon E, Coutant C, Doimi FD, Ibrahim NK, Gong Y, Hortobagyi GN, Hess KR, Symmans WF, Pusztai L: **Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer.** *Clin Cancer Res* 2010, **16:**5351-5361.
96. Potti A, Dressman HK, Bild A, Riedel RF, Chan G, Sayer R, Cragun J, Cottrill H, Kelley MJ, Petersen R, Harpole D, Marks J, Berchuck A, Ginsburg GS, Febbo P, Lancaster J, Nevins JR: **Genomic signatures to guide the use of chemotherapeutics.** *Nat Med* 2006, **12:**1294-1300.
97. Bonnefoi H, Potti A, Delorenzi M, Mauriac L, Campone M, Tubiana-Hulin M, Petit T, Rouanet P, Jassem J, Blot E, Becette V, Farmer P, Andre S, Acharya CR, Mukherjee S, Cameron D, Bergh J, Nevins JR, Iggo RD: **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:**1071-1078.
98. Coombes KR, Wang J, Baggerly KA: **Microarrays: retracing steps.** *Nat Med* 2007, **13:**1276-1277; author reply 1277-1278.
99. Liedtke C, Wang J, Tordai A, Symmans WF, Hortobagyi GN, Kiesel L, Hess K, Baggerly KA, Coombes KR, Pusztai L: **Clinical evaluation of chemotherapy response predictors developed from breast cancer cell lines.** *Breast Cancer Res Treat* 2010, **121:**301-309.

100. Baggerly KA, Coombes KR, Neeley ES: **Run batch effects potentially compromise the usefulness of genomic signatures for ovarian cancer.** *J Clin Oncol* 2008, **26:**1186-1187; author reply 1187-1188.
101. Collingridge D: **Expression of concern--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* 2010, **11:**813-814.
102. Borst P, Wessels L: **Do predictive signatures really predict response to cancer chemotherapy?** *Cell Cycle* 2010, **9:**4836-4840.
103. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, Andre S, Piccart M, Campone M, Brain E, Macgrogan G, Petit T, Jassem J, Bibeau F, Blot E, Bogaerts J, Aguet M, Bergh J, Iggo R, Delorenzi M: **A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer.** *Nat Med* 2009, **15:**68-74.
104. Ashworth A, Lord CJ, Reis-Filho JS: **Genetic interactions in cancer progression and treatment.** *Cell* 2011, **145:**30-38.
105. Juul N, Szallasi Z, Eklund AC, Li Q, Burrell RA, Gerlinger M, Valero V, Andreopoulou E, Esteva FJ, Symmans WF, Desmedt C, Haibe-Kains B, Sotiriou C, Pusztai L, Swanton C: **Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials.** *Lancet Oncol* 2010, **11:**358-365.
106. Li Y, Zou L, Li Q, Haibe-Kains B, Tian R, Li Y, Desmedt C, Sotiriou C, Szallasi Z, Iglehart JD, Richardson AL, Wang ZC: **Amplification of LAPTM4B and YWHAZ contributes to chemotherapy resistance and recurrence of breast cancer.** *Nat Med* 2010, **16:**214-218.
107. Vollebergh MA, Lips EH, Nederlof PM, Wessels LF, Schmidt MK, van Beers EH, Cornelissen S, Holtkamp M, Froklage FE, de Vries EG, Schrama JG, Wesseling J, van de Vijver MJ, van Tinteren H, de Bruin M, Hauptmann M, Rodenhuis S, Linn SC: **An aCGH classifier derived from BRCA1-mutated breast cancer and benefit of high-dose platinum-based chemotherapy in HER2-negative breast cancer patients.** *Ann Oncol* 2010, Dec 6. [Epub ahead of print].
108. Bauer JA, Chakravarthy AB, Rosenbluth JM, Mi D, Seeley EH, De Matos Granja-Ingram N, Olivares MG, Kelley MC, Mayer IA, Meszoely IM, Means-Powell JA, Johnson KN, Tsai CJ, Ayers GD, Sanders ME, Schneider RJ, Formenti SC, Caprioli RM, Pietenpol JA: **Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation.** *Clin Cancer Res* 2010, **16:**681-690.
109. Munkacsy G, Abdul-Ghani R, Mihaly Z, Tegze B, Tchernitsa O, Surowiak P, Schafer R, Gyorffy B: **PSMB7 is associated with anthracycline resistance and is a prognostic biomarker in breast cancer.** *Br J Cancer* 2010, **102:**361-368.
110. Jansen MP, Foekens JA, van Staveren IL, Dirkzwager-Kiel MM, Ritstier K, Look MP, Meijer-van Gelder ME, Sieuwerts AM, Portengen H, Dorssers LC, Klijn JG, Berns EM: **Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling.** *J Clin Oncol* 2005, **23:**732-740.
111. Oh DS, Troester MA, Usary J, Hu Z, He X, Fan C, Wu J, Carey LA, Perou CM: **Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers.** *J Clin Oncol* 2006, **24:**1656-1664.
112. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, Gillet C, Ellis P, Ryder K, Reid JF, Daidone MG, Pierotti MA, Berns EM, Jansen MP, Foekens JA, Delorenzi M, Bontempi G, Piccart MJ, Sotiriou C: **Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen.** *BMC Genomics* 2008, **9:**239.
113. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, Daxenbichler G, Desmedt C, Domont J, Marth C, Delaloge S, Bauernhofer T, Valero V, Booser DJ, Hortobagyi GN, Pusztai L: **Genomic index of sensitivity to endocrine therapy for breast cancer.** *J Clin Oncol* 2010, **28:**4111-4119.
114. Loi S, Haibe-Kains B, Majjaj S, Lallemand F, Durbecq V, Larsimont D, Gonzalez-Angulo AM, Pusztai L, Symmans WF, Bardelli A, Ellis P, Tutt AN, Gillett CE, Hennessy BT, Mills GB, Phillips WA, Piccart MJ, Speed TP, McArthur GA, Sotiriou C: **PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer.** *Proc Natl Acad Sci U S A* 2010, **107:**10208-10213.
115. McDermott U, Downing JR, Stratton MR: **Genomics and the continuum of cancer care.** *N Engl J Med* 2011, **364:**340-350.
116. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC, O'Connell P: **Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer.** *Lancet* 2003, **362:**362-369.

doi:10.1186/bcr2890
**Cite this article as**: Colombo PE, *et al.*: Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction. *Breast Cancer Research* 2011, **13:**212.