REVIEW ARTICLE

# Prediction and prognosis: impact of gene expression profiling in personalized treatment of breast cancer patients

Michael R. Mallmann · Andrea Staratschek-Jox ·
Christian Rudlowski · Michael Braun · Andrea Gaarz ·
Matthias Wolfgarten · Walther Kuhn · Joachim L. Schultze

**Abstract** Breast cancer is a complex disease, whose heterogeneity is increasingly recognized. Despite considerable improvement in breast cancer treatment and survival, a significant proportion of patients seems to be over- or undertreated. To date, single clinicopathological parameters show limited success in predicting the likelihood of survival or response to endocrine therapy and chemotherapy. Consequently, new gene expression based prognostic and predictive tests are emerging that promise an improvement in predicting survival and therapy response. Initial evidence has emerged that this leads to allocation of fewer patients into high-risk groups allowing a reduction of chemotherapy treatment. Moreover, pattern-based approaches have also been developed to predict response to endocrine therapy or particular chemotherapy regimens. Irrespective of current pitfalls such as lack of validation and standardization, these pattern-based biomarkers will prove useful for clinical decision making in the near future, especially if more patients get access to this form of personalized medicine.

M. R. Mallmann (✉) · C. Rudlowski · M. Braun ·
M. Wolfgarten · W. Kuhn
Department of Obstetrics & Gynecology, Center for Integrated Oncology, University Hospital of Bonn,
Sigmund-Freud-Strasse 25,
53105 Bonn, Germany
e-mail: michael.mallmann@ukb.uni-bonn.de

M. R. Mallmann · A. Staratschek-Jox · A. Gaarz · J. L. Schultze
LIMES (Life and Medical Sciences Bonn) Institute,
Genomics and Immunoregulation, University Bonn,
Carl-Troll-Strasse 31,
53115 Bonn, Germany

## Introduction

Breast cancer is the most common cancer of women in the western world. The overall death rate has been significantly reduced in the last decades, but depending on subtype and stage, still a significant portion of patients will suffer from relapse or even die of the disease [1, 2]. While up to 70% of patients with breast cancer can be cured nowadays, a significant proportion of these patients is overtreated. It remains a challenge to identify those patients who will indeed profit from current treatment strategies and also to develop innovative concepts for patients currently at high-risk for relapse after treatment. For this reason, the identification of reliable prognostic biomarkers together with the development of clinically efficient therapies is urgently needed [3]. Today, the prognostic clustering of breast cancer in daily routine relies on the determination of a limited set of molecular markers (e.g. estrogen receptor (ER), progesterone receptor (PR) and epidermal-growth-factor receptor 2 (HER2, also referred to as Her2/neu, ErbB-2)) mostly by semi-quantitative assays e.g. by immunohistochemistry (Fig. 1). Clearly, some of these markers are first examples of personalized medicine and targeted treatment since for instance only the determination of ER-expression by immunohistochemistry allows for a directed anti-hormonal therapy with receptor blockade or inhibition, or both. [4]. Moreover HER2-overexpression has paved the way for anti-HER2 treatment with the humanized monoclonal antibody trastuzumab [5–7] or the small-molecule inhibitor of the tyrosine kinase domains of HER1 and HER2, lapatinib [8–10]. The best HER2-targeted treatment option together with chemotherapy in patients with metastasized but operable breast cancer is currently assessed in clinical trials [11].

In addition to tissue based markers that have prognostic and predictive value, blood-based proteomic tests for early
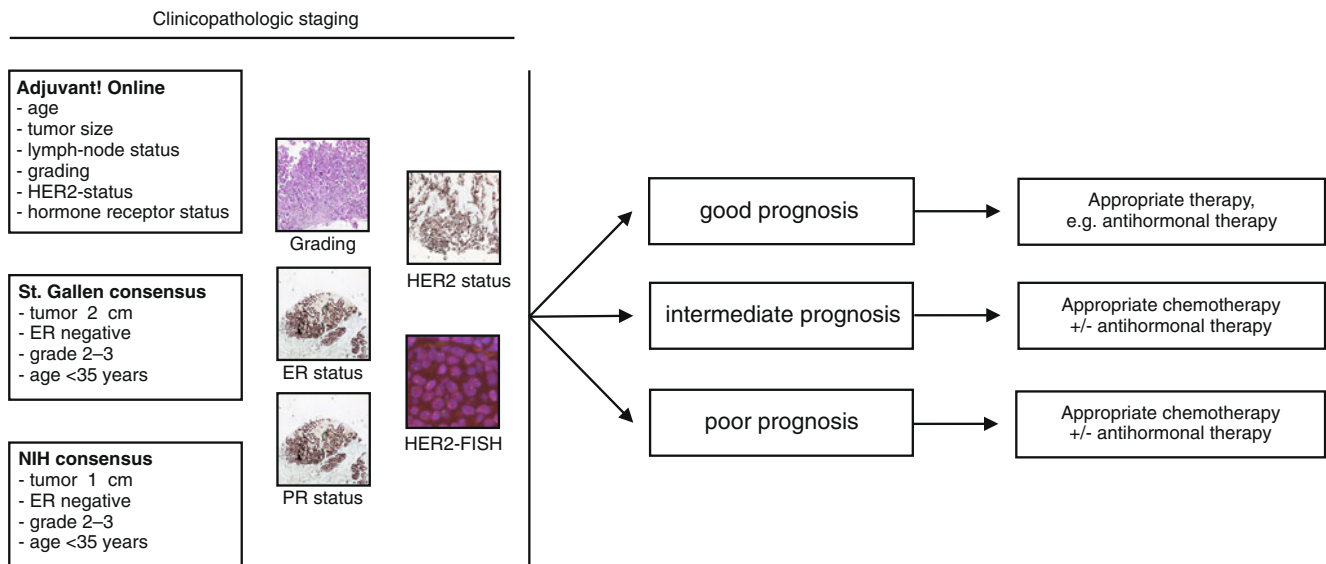
**Fig. 1** Current clinocopathologic decision making. Patients are currently allocated into clinical risk groups by several mechanisms. Clinical parameters such as tumor size, lymph-node status and age as well as pathologic parameters such as histologic grading, hormone receptor status and HER2-status are main factors for risk assignment in breast cancer therapy. This risk assignment results in allocation into a low risk group that may be properly treated with hormonal therapy only or other treatments and a high risk group mainly treated with chemotherapy if no patient specific contradictions apply (e.g. waiving of anthracycline-based chemotherapy in patients with existing heart failure). The intermediate risk group due to uncertain outcome is mainly treated with chemotherapy the best choice of therapy currently under intense clinical studies

detection of breast cancer are emerging. Consequently, non-invasive diagnostic approaches based on pathology-specific molecular-patterns in blood might identify breast cancer in an earlier phase of their disease [12–14] and might be used to easily monitor therapy responses [15].

Nonetheless, breast cancer is clinically heterogeneous with varying response to treatment, even when taking into account the above mentioned therapeutic targets. The established methods that are suited to study one gene at a time do not seem to have the power to cover this clinical heterogeneity, which is likely to be due to a complex set of multiple somatic mutations, epigenetic changes and genomic rearrangements [16, 17]. To overcome the limitation of single gene or protein biomarkers, the implementation of DNA microarray technology nearly a decade ago has enabled the quantitative measurement of complex gene expression-patterns ("gene expression profiling") in breast and other cancers and has paved the way to new pattern-based biomarker strategies.

DNA array technology has been successfully used to identify subtypes in breast cancer based on their specific gene expression patterns [18]. In general, a molecular taxonomy that allocates breast cancer samples into at least five subtypes, termed basal-like, ErbB2, luminal A, luminal B and normal like breast cancer, has been reproduced by several independent groups and is generally accepted as gene-signature based molecular classification [19–22]. Interestingly, these molecular patterns seem to be remark-ably stable between primary tumor and distant metastases [23].

**Prognostic gene profiles**

The identification of pattern-based biomarkers for prognosis is a major field of current clinical research in many cancer types including breast cancer. Today, the majority of patients with early breast cancer will receive adjuvant chemotherapy. Only a minority are likely to benefit from such therapy, but all of them will be affected by its toxicity [24]. Consequently, the identification of prognostic markers identifying the subset of patients eligible for a "watchful waiting" procedure and/ or adjuvant anti-hormonal/ anti-HER2 therapy could help to minimize therapy-induced side effects. Consecutively, expression-based outcome prediction by use of prognostic signatures has been explored in a variety of studies.

The first breast cancer prognostic signature to be described has been a 70-gene signature by van 't Veer et al. [25]. It was developed based on the analysis of 78 young (<55 years) patients with sporadic lymph-node-negative stage I or II breast cancer that were followed up for the development of distant metastasis during 5 years after diagnosis. The retrospective allocation to the 'good'—metastasis free—prognosis group or 'poor' prognosis group developing metastasis solely relied on the 70-gene signature

and was irrespective of ER, PR or HER2neu expression, well defined markers analyzed routinely in newly diagnosed breast cancers. To validate the 70-gene signature in an independent validation cohort, the authors studied a validation cohort of 295 consecutive patients with lymph-node-negative and lymph-node-positive breast cancer (including 61 of the 78 patients with lymph-node-negative disease who were involved in the previous study) [26]. In this cohort, the 70 gene prognosis profile was a strong predictor of the development of distant metastasis in patients with both lymph-node-negative as well as lymph-node-positive disease. Interestingly, patients with lymph-node-negative disease and those with lymph-node-positive disease were evenly distributed in the 'good' and the 'poor' prognosis group, indicating that the 70-gene prognosis profile might be independent of lymph-node status [26]. The authors then compared the probability that patients classified according to either the 70-gene expression profile, to St. Gallen criteria [27], or to the National Institutes of Health (NIH) consensus criteria [28] would remain free of distant metastasis. They showed that the 70-gene prognosis-profile assigns more patients with lymph-node-negative disease to a low-risk ('good' prognosis signature) group and that these low-risk patients had a higher likelihood of metastasis free survival than those classified according to the traditional methods [26]. A tendency towards a higher rate of distant metastasis was observed for the 'poor' prognosis signature group as compared to high-risk patients identified by St. Gallen or NIH criteria.

In an independent validation study conducted by the TRANSBIG consortium, the clinical utility of the 70 gene signature was further assessed in 302 patients with early-stage ER- negative and ER-positive breast cancer [29]. The 70-gene prognosis signature was a statistically significant prognostic factor for time to distant metastasis (the outcome the signature had been developed on) and overall survival. 90% of women in the 'good' prognosis group and 71% within the 'poor' prognosis group remained free from distant metastasis at 10 years of follow up, respectively. A second validation study confirmed the data showing that at 5 years, the probability of remaining free of distant metastasis was 98% for the 'good' and 78% for the 'poor' prognosis signature patients in an independent set of 123 patients [30]. As lymph-node status is one of the most important prognostic factors in breast cancer, a study on 241 patients extended the 70-gene signature prognostic value from lymph-node negative to lymph-node positive breast cancer. Patients with 1–3 positive axillary lymph nodes have a 10- year distant metastasis-free survival of 91% and 76% in the 'good' and the 'poor' prognosis group, respectively [31]. The 70-gene signature classified 41% of patients among lymph-node positive patients into a low-risk

group for metastasis. With common clinicopathological parameters these patients would have been scored as high-risk patients for recurrence and it is this group of patients that might greatly benefit from reducing chemotherapy [31]. As there might be a difference in gene expression between premenopausal and postmenopausal women [32], two studies investigated the prognostic value of the 70-gene signature in postmenopausal women providing evidence that the 70-gene prognosis signature is also well suited to predict metastasis at 5 years in the postmenopausal woman aged 55 to 70 years [33]. In this patient group, individuals belonging to the 'good' prognosis group had a 93% and 99%, compared to a 72% and 80% 5-year distant metastasis-free and breast-cancer specific survival, respectively [33]. As in the preceding studies, in older patients the 70-gene profile illustrates a high power in respect to its negative predicting value, further validated in an independent study of 100 older lymph-node negative patients [34]. It is important to mention that distant metastasis after 5 years is also significantly less accurately predicted by the signature in older patient cohorts [33]. This is important to know when interpreting the data, as 25% of all metastases occur more than 5 years after initial diagnosis, and the percentage might be higher in the subset of patients with ER-positive breast cancer, the subtype mostly found in older patient [35]. To facilitate the use of the 70-gene prognosis profiler as diagnostic test, the 70-gene profile was translated into a customized microarray with only 1,600 instead of 25,000 probes, merchandised as MammaPrint® (Agendia, Huntington Beach, CA), which has been cleared by the Food and Drug Administration (FDA) in 2007 [36]. Statistical analysis has shown that the 70-gene signature not only correlates with established factors of prognosis like age, grading and ER-status, but outperforms the well-established prognostic algorithms such as the St. Gallen criteria [26, 29, 30, 37]. The 70-gene-profile reliably identifies high-risk patients that require chemotherapy. Even more important, the 70-gene-profile identifies patients with low risk of recurrence, who could spare chemotherapy. Patients that belong to a low-risk-group according to the 70-gene signature shared a better recurrence free survival than patients, who were classified according to St. Gallen-criteria or NIH-criteria. It is important to mention that this signature mainly discriminates high-risk and low-risk situations in patients with ER-positive disease. In patients with ER-negative disease, the predictive power is limited. All these studies have been done retrospectively and additionally, the differences in risk of metastasis between the clinicopathological guidelines and the prognosis signature have been tremendous with up to 39% different allocation in both directions [38]. Consequently, in 2007 the TRANSBIG Research Consortium initiated a large prospective, multicenter, controlled randomized MINDACT (Microarray in Node

negative Disease may Avoid ChemoTherapy) trial (Fig. 2) [39]. In this study 6000 nodal-negative patients are included and the recurrence risk is calculated on the basis of either the 70-gene prognosis score signature or six classical factors of prognosis (age, size, lymph-node-status, grading, HER2-status, hormone receptor status). If both tests allocate the patient to a high-risk group, patients are randomized into an anthracyclin-based chemotherapy or a Docetaxel/ Gemcitabine chemotherapy, whereas if both tests allocate the patient to a low-risk group, no further chemotherapy is recommended. In both arm, patients whose tumors show ER-expression are randomized to either letrozol or tamoxifen followed by letrozol. It is expected that for 1/3 of the patients, the tests show different results with a low-risk genomic profile and a high-risk classical profile are vice versa. In these cases the primary objective of the MINDACT trial is to confirm that patients with a "good" molecular prognosis score but "high-risk" clinical prognosis score be safely spared chemotherapy without affecting distant-metastasis free survival. In a first treatment decision randomization patients are therefore randomized in either chemotherapy or no chemotherapy. Patients assigned to chemotherapy are randomized into an anthracyclin-based chemotherapy or Docetaxel/ Gemcitabine and patients whose tumors show ER-expression undergo a third randomization into either letrozol or tamoxifen followed by letrozol treatment.

Another pattern-based biomarker assay using the expression levels of 21 genes assessed by RT-PCR from paraffin-embedded tumor tissue is the Oncotype DX assay (Genomic Health, Redwood City, CA) [40]. Two studies were conducted to evaluate the Oncotype DX 21-gene assay in terms of predicting recurrence-free survival in lymph-node negative breast cancer without adjuvant therapy. In 149 hormonal and chemotherapy naive patients no clear association between the RT-PCR based recurrence score (RS) and distant recurrence was found (18%, 38% and 28% rates of distant recurrence at 10 years in the low, intermediate- and high-risk group, respectively) [41]. A case-control-study including 790 lymph-node negative breast cancer patients not treated with adjuvant chemotherapy, however, showed a significant association between RS and breast cancer death with a risk of 6.2%, 17.8% and 19.9% for breast cancer death in the low-, intermediate- and high-risk in the ER-positive patient group [42]. The advantage of the Oncotype DX tests is the expression analysis with RT-PCR on formalin-fixed tissue which does not rely on fresh frozen tumor samples.

In a complementary analysis to the 70-gene prognosis score, the Rotterdam 76-gene signature was obtained analyzing 286 patients with lymph-node-negative disease who had not received adjuvant systemic treatment [43]. Using this 76-gene signature, 93% of women with a good

prognosis signature will survive without distant-metastasis after 60 months and 88% after 80 months, respectively. In two subsequent studies, these results were validated. In the first study on 180 stage I-II breast cancer patients, patients with a good prognosis score showed 5- and 10-year distant metastasis-free survival of 96% and 94%, and those with a poor prognosis group 74% and 65%, respectively [44]. The positive prediction value (PPV) and negative prediction value (NPV) were 38% and 94%, respectively. The second study with 198 patients showed 5- and 10-year distant metastasis-free survival of 98% and 94%, and those with a poor prognosis group 76% and 73%, respectively [45].

As none of the aforementioned studies has concentrated on the group of patients overexpressing HER2, two groups delineated the heterogeneity of HER2 positive breast cancer. The HER gene is amplified or overexpressed in 15–20% of patients and these patients are mostly classified in the high-risk group by both the 70-gene signature and the 21-gene recurrence score [46, 47]. Using principle component analysis (PCA), Alexe et al. identified 105 lymphocyte-associated genes in two core subgroups denoted $HER2+_I$ and $HER2+_{NI}$ in HER2-overexpressing node-negative patients [48]. Patients with a $HER2+_I$ signature showed a moderate to marked lymphocyte infiltrate with a recurrence rate of 11%, whereas patients with a $HER2+_{NI}$ signature showed only minimal lymphocytic infiltrate and a recurrence rate of 58%. In this study, immune cell infiltration was associated with lower recurrence rate which might suggest that in these cases the tumor is more effectively recognized and eliminated by the immune system than in cases with low immune cell infiltration.

In 58 HER2-amplified trastuzumab untreated tumors, Staaf and collegues established a 158-gene HER2-derived prognostic predictor (HDPP) that classifies HER2-overexpressing tumors in a 'good' and a 'poor' prognosis group in both overall and distant-metastasis-free survival [47]. In multivariate analysis, the HDPP signature was associated with an improved stratification into worse overall survival and distant-metastasis free survival and, if compared to the HER2 prognostic predictor developed by Alexe et al., showed a similar categorization of HER2-overexpressing tumors into a 'good' and a 'poor' prognosis group with the exception of one patient. Importantly, the prognostic profile performs superior on the group of HER2-overexpressing tumors than MammaPrint® or the Oncotype DX test, but is not prognostic in luminal A, luminal B or normal-like-tumors. Although this data was promising, the predictive power of the HDPP signature in trastuzumab treatment, the standard care today for patients with HER2-overexpressing breast cancer, was only investigated in a small set of 22 patients treated neoadjuvantly with
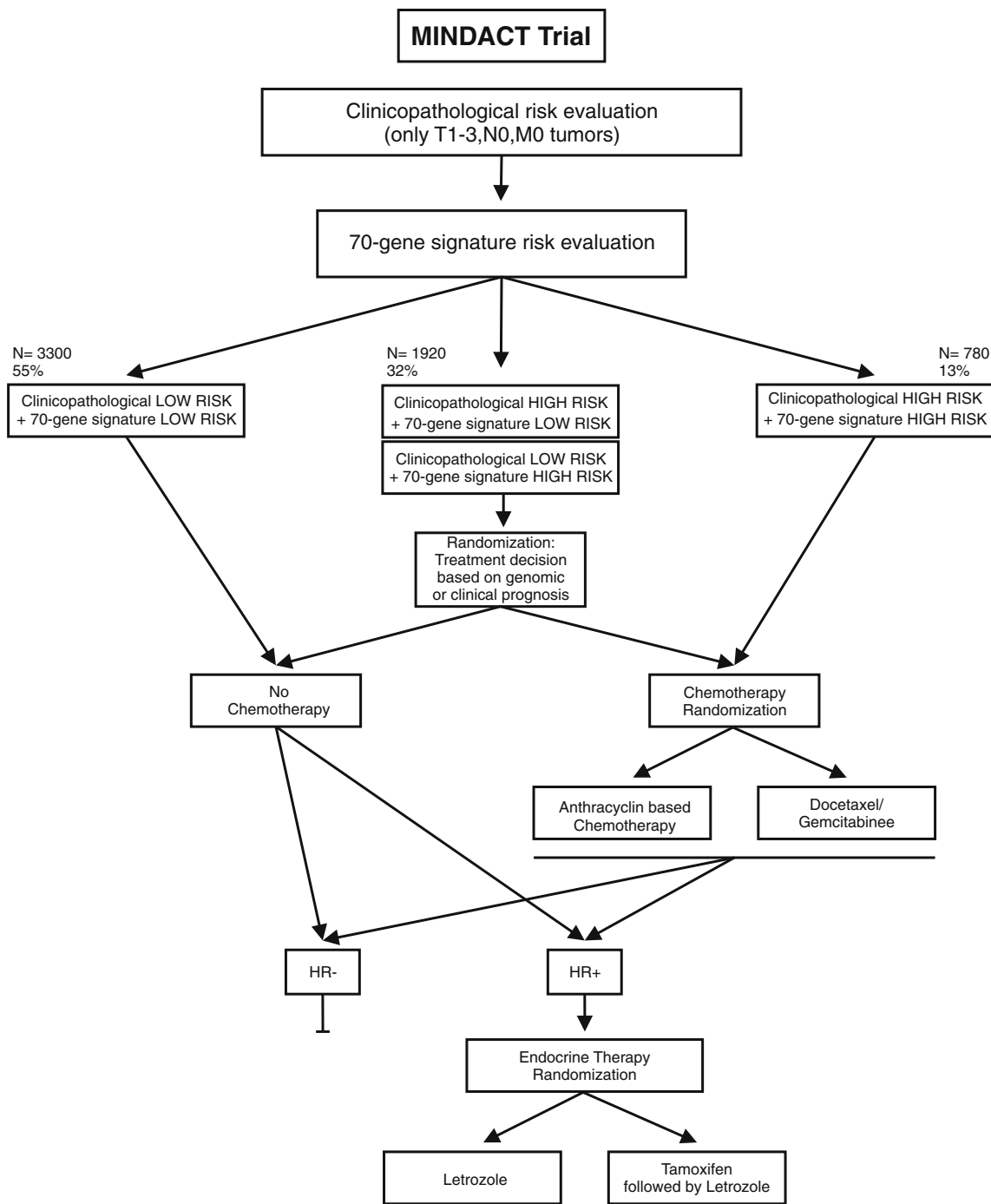
**Fig. 2** Overview of the MINDACT Trial. The MINDACT trial compares the 70-gene expression signature with the common clinical-pathological prognostic tool Adjuvant! Online in selecting patients for adjuvant chemotherapy in LN-negative but both ER-positive and ER-negative breast cancer. Patients with both a low-risk profile in Adjuvant! Online and the 70-expression profile are not treated with chemotherapy. If their tumors show ER-expression, patients are randomized to either letrozol or tamoxifen followed by letrozol. Patients with both high-risk profiles in Adjuvant! Online and the 70-expression profile are randomized into an anthracyclin-based chemotherapy or a Docetaxel/ Gemcitabine chemotherapy. Again, if their tumors show ER-expression, patients undergo a second randomization into either letrozol or tamoxifen followed by letrozol. The primary study population, patients with either a high-risk profile in the Adjuvant! Online tool but low-risk 70-expression profile or vice versa can be randomized into different treatment arms. In a first treatment decision randomization, the patients are randomized in either chemotherapy or no chemotherapy. Patients assigned to chemotherapy are randomized into an anthracyclin-based chemotherapy or Docetaxel/ Gemcitabine and patients whose tumors show ER-expression undergo a third randomization into either letrozol or tamoxifen followed by letrozol treatment

trastuzumab and vinorelbine, to small to small to draw definitive conclusion.

In 2003, Huang et al developed a gene expression profile that classified breast tumors by their likelihood of having concomitant lymph-node metastasis at the time of diagnosis and their 3-year recurrence risk. Interestingly, a comparison of this gene set, which predicts the development of local metastasis to the regional lymph-nodes, to gene sets that predict development of distant metastasis show different genetic fingerprints suggesting local and distant metastasis to display different molecular programs [26, 49].

Alternative approaches to gene expression classifiers based on outcome emerged that base on the assumed underlying biological mechanisms ("hypothesis-driven" or "bottom-up" approach) instead of outcome-related signatures ("top-down approach"). The "Wound response gene expression signature" was additionally applied to the gene expression data from the 295 breast cancer patients initially used to establish the 70-gene prognostic profile [50]. A set of "core serum response" genes was used that had been previously identified in vitro in fibroblasts activated with serum and comprised processes in wound healing such as matrix remodeling and angiogenesis [51]. With a threshold set to identify 90% of patients with subsequent metastasis, patients were either stratified in a high-risk group with "activated" or in a low-risk group with "quiescent" wound response signature, respectively. As does the 70-gene prognosis signature, the wound response signature provides independent prognostic information in multivariate analysis. In a set of lymph-node negative (51%) and lymph-node positive (49%) breast cancer patients, distant-metastasis-free survival at 10 years was 51% in the "activated" wound-response signature group, 75% in the "quiescent" wound-response signature negative group, and overall survival was 50% and 84%, respectively, indicating the wound-response signature to be a powerful prognostic marker. Comparison to the traditional clinical NIH and St. Gallen consensus criteria revealed that the wound-response signature would add information for patients that have been stratified as high-risk by NIH or St. Gallen criteria, since this test would have spared 30% (with high-risk NIH and St. Gallen criteria but quiescent wound response signature) from chemotherapy that otherwise are assigned to chemotherapy [50]. When comparing the wound response signature to the 70-gene signature, the statistically defined 70-gene signature outperformed the biological derived wound-response signature in terms of sensitivity (85.2% and 59.1%, respectively), but not specificity (49.3% and 64.3%). It is not surprising that the 70-gene signature outperforms the biological derived signature, as other factors intrinsic to the tumor and not reflected by a wound response gene signature are more likely been represented by a statistically defined than by an hypothesis-driven gene

signature. In an alternative approach than to compare the wound-response signature to the 70-gene signature, both the 70-gene and the wound response signatures were incorporated into a common decision tree to optimize risk stratification. ER-receptor positive lymph-node positive patients were primarily stratified using the 70-gene prognosis signature into 'good' and 'poor' prognosis groups. The good prognosis group (38% of patients), henceforward termed "very good prognosis group", had a distant-metastasis-free survival of 89% at 10 years. The poor-prognosis group can be subsequently divided into a quiescent-wound-response subset (22% of patients) with a distant-metastasis-free survival of 78% and an activated-wound-response subset (40% of patients) with a distant-metastasis-free survival of 47% at 10 years, respectively [50].

To better discriminate breast cancer samples that are classified as histologic grade 2, which is barely informative for clinical decision making, a 97-gene expression signature was developed that, when applied to histologic grade 2 breast cancer samples, allocates the tumor into two prognostic different subgroups namely 'genomic grade 1' and 'genomic grade 3' cancer [52]. As genomic grade represents a prognostic factor with regard to cancer aggressiveness, the 97-gene expression signature was found to yield similar results in terms of predicting distant free survival compared to the 70-gene classifier/signature.

A biologic hypothesis driven assay investigated the role of breast cancer stem cell associated genes. By comparing the gene expression profile of putative $CD44^+CD24^-$ breast cancer stem cells with normal breast epithelial cells, a gene expression profile of 186 genes were found to be associated with breast cancer stem cells [53]. This "invasiveness gene signature" (IGS) stratifies high-risk lymph-node negative early breast cancer into a 'good' and a 'poor' prognosis group of patients with a 10-year rate of metastasis-free survival of 81% and 57%, respectively.

Taken together, the new set of gene expression based profilers to predict prognosis show an exciting high prognostic power. Although they show different power among each other to predict prognosis of untreated patients, they are of equal value or even outcompete traditional decision criteria like Adjuvant! Online, the St. Gallen consensus or the NIH consensus criteria. With these new prognostic markers, in the near future stratification of patients into risk groups according to the gene expression profile of their tumors might be feasible (Fig. 3). The former high number of patients with intermediate-risk tumors might be allocated to either the gene expression profile high-risk group or the gene expression profile low-risk group. Patients in the gene expression profile high-risk group might undergo chemotherapy and hormonal therapy if ER-receptor positive. Patients in the gene expression profile low-risk group might be objected to hormonal therapy, minimizing unnecessary chemotherapy.

**Fig. 3** Gene expression based prognostic profiling. Among all patients with breast cancer, patients with a "good prognosis" and a "poor prognosis" gene expression profile are identified and are allocated into risk groups according to the gene expression profile of their tumors. Immunohistochemistry might either still be needed to allocate patients to antihormonal and anti-HER2 treatment or might be replaced by gene expression based profiles of ER- expression, HER2-overexpression and other target profiles. In this setting, the former intermediate risk group is allocated to either the gene expression profile high-risk or the gene expression profile low-risk group and unnecessary chemotherapy is minimized



Prognosis using pattern-based biomarkers might therefore spare many patients with a low risk of recurrence chemotherapy that they otherwise would have assigned to chemotherapy (Fig. 4).

### Prediction—response to endocrine therapy

In addition to questions concerning prognosis several attempts have been undertaken to develop predictive pattern-based biomarkers for patients with breast cancer. A large proportion of breast cancers show estrogen dependent growth. Already a long time ago oophorectomy has been shown to cause regression of breast cancer and today, estrogen deprivation remains a key therapeutic approach to treat breast cancer with ER-expression [54]. Standard adjuvant endocrine therapy has long consisted of a treatment with tamoxifen for 5 years, which improved both disease-free as well as overall survival. However, the likelihood for patients with ER-positive breast cancer to develop distant metastasis after surgery and adjuvant tamoxifen alone is still 15% after 10 years [55]. Consequently, the majority of patients would be overtreated if chemotherapy would be administered to everyone. Today, absolute levels of ER and PR expression as positive and HER2 and EGFR as negative predictors remain the best predictors of tamoxifen response [56, 57]. Antiestrogen therapy, when compared to chemotherapy, is well tolerated and only associated with minor toxicities [58]. However, failure to respond or to develop early resistance to tamoxifen is seen in approximately 25% of $ER^+/PR^+$, 66% of $ER^+/PR^-$ and 55% of $ER^-/PR^+$ breast tumors [58]. Particularly in these patients the use of alternative endocrine therapies or chemotherapy is indicated. However, to reduce treatment-associated side-effects and increase therapy efficacy, prediction of tamoxifen treatment response particularly in ER-positive early stage breast cancer would be required. Addressing this important clinical question, several groups have started to develop gene expression profiles to predict tamoxifen treatment outcome.

Paik and colleagues selected a candidate-gene approach to develop a gene expression signature that predicts the likelihood of distant recurrence in patients with lymph-
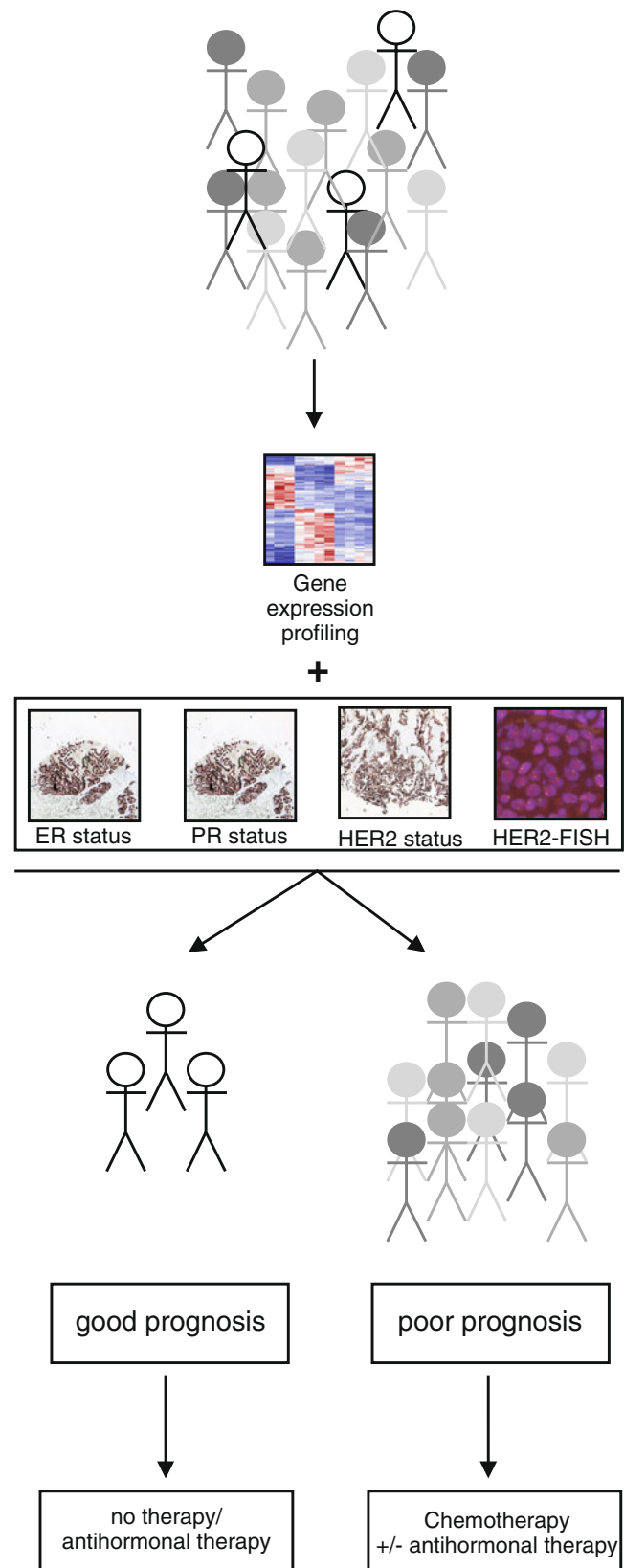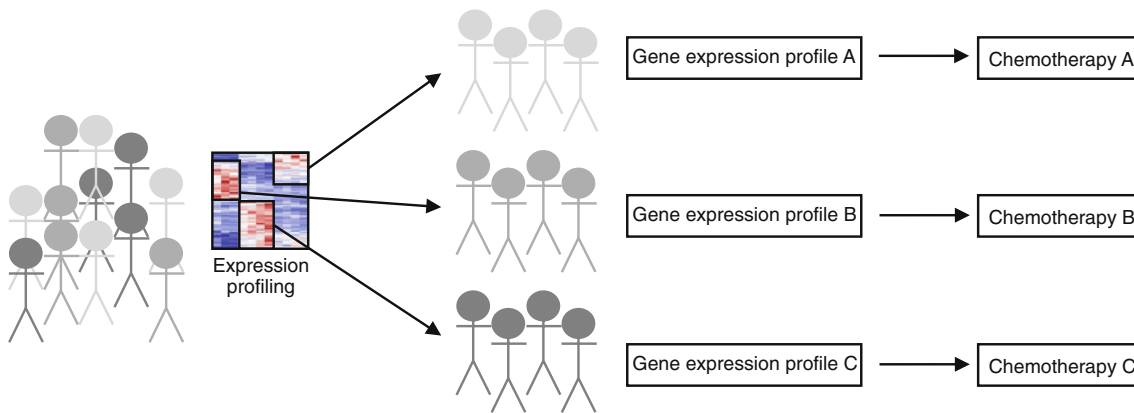
**Fig. 4** Predictive gene expression based assignment to chemotherapy. Patients which harbor a high-risk gene expression profile according to Fig. 3 and which can undergo chemotherapy according to clinical risk management (e.g. age, comorbidities, ECOG status) might be further stratified. By extraction of gene expression profilers predictive for specific chemotherapies, patients are allocated to the treatment they profit most of

node-negative, ER-positive breast cancer that had been treated with tamoxifen [40]. This assay, later established as Oncotype DX (Genomic Health, Redwood City, CA), combines a panel of 21 genes, which had been selected from preliminary studies. The expression of 16 genes grouped on the basis of function and normalized to the expression of five reference genes builds a quantitative continuous recurrence score (RS) and this score estimates the probability of recurrence after 10 years. By grouping patients according to the rate of distant metastasis, low-risk (less than 18) and high-risk (31 or higher) groups were generated on the basis of data from the NSABP B-14 trial. In this trial, 6.8% of patients in the low-risk group and 30.5% of patients in the high-risk group developed distant metastasis at 10 years after diagnosis, respectively. The probability for recurrence in the high-risk group is therefore in the range observed for lymph-node-positive breast cancer. In several studies, the risk for patients with low recurrence score and node-negative breast cancer treated with tamoxifen is below 10% at 10 years [59]. One of these studies showed that patients with a high recurrence score might benefit from adjuvant CMF chemotherapy (cyclo-phosphamide, methotrexate, and fluorouracil) [59]. Since the clinical trials leading to inclusion of the RS into the decision tree for therapy by the American Society of Clinical Oncology (ASCO) included only patients with node-negative breast cancer, the predictive value of the 21-gene RS was extended to postmenopausal node-positive breast cancer in the Southwest Oncology Group (SWOG)-8814, INT-0100 trial [60]. In this trial, the RS was highly prognostic for disease-free survival within the tamoxifen group. The 10-year disease-free survival were 60%, 49%, and 43% for the low, intermediate, and high-risk groups, respectively, with an overall hazard ratio of 2.64 (95% CI 1.33–5.27; $p=0.006$) for a 50-point difference in the continuous RS. The hazard ratio for a 50-point difference

was 5.55 (95% CI 2.32–3.28; $p=0.0002$) in the first 5 years. Interestingly, the RS was not prognostic any more beyond the first 5 years. However, the RS showed prognostic value for overall survival at 10 years for patients treated with tamoxifen alone. The 10-year overall survival were 77%, 68%, and 51% for the low, intermediate, and high-risk groups, respectively, with an overall hazard ratio of 4.42 (95% CI 1.96–9.97; $p=0.0006$) for a 50-point difference in the continuous RS. In this study, the RS also showed value in predicting the efficacy of the more currently used anthracycline-based chemotherapy regimen CAF (cyclo-phosphamide, doxorubicin, and fluorouracil) in patients with node-positive ER-positive breast cancer, as patients with a high recurrence score ($\geq 31$) showed a benefit from adjuvant chemotherapy (HR 0.59, 95% CI 0.35–1.01; $p=0.033$) in the first 5 years, although there was no additional prediction beyond 5 years [61]. No benefit was observed for patients with a RS of less than 18 (HR 1.02, 95% CI 0.54–1.93; $p=0.97$) or with a RS between 18 and 30 (HR 0.72, 0.39–1.31; $p=0.48$). The 10-year disease-free survival for patients with a low RS were 64% in the CAF-T group and 60% in the tamoxifen-alone group, and for patients with a high recurrence score, 55% and 43%, respectively.

Since aromatase inhibitors have been established in clinical practice for treatment of ER-positive breast cancer, the performance of the RS for distant recurrence in postmenopausal women with localized N0 and node-positive breast cancer was evaluated in the TransATAC study (The Arimidex, Tamoxifen, Alone or in Combination) [62]. For patients with lymph-node negative ER-positive breast cancer, the rates of distant recurrence at 9 years were 4% (95% CI, 3%–7%), 12% (95% CI, 8%–18%), and 25% (95% CI, 17%–34%). For patients with lymph-node positive ER-positive breast cancer, the rates of distant recurrence at 9 years were 17% (95% CI, 12%–24%), 28% (95% CI, 20%–39%), and 49% (95% CI, 35%–

64%) [62]. In summary, the 21-gene recurrence score adds important information to the traditional pathological approaches, but some problems remain. The Oncotype DX assay is, although prognostic and predictive before 5 years, not prognostic and predictive after 5 years. Furthermore, the test is expensive and there is still no data on the predictive value of the assay on modern taxane-based chemotherapy [63]. Nonetheless, the 21-gene RS is already in daily clinical use and supports changes in treatment decisions by both medical oncologists but also patients themselves [64].

Several other biomarkers are currently in development, often in an attempt to reduce the number of data points to be assessed thereby potentially also reducing costs. Along these lines, Ma et al. developed a two-gene expression ratio from 60 laser-capture microdissected breast cancer tissue [57]. This assay predicts response to tamoxifen by the ratio of HOXB13:IL17BR with both high HOXB13 (homeo domain-containing protein) and low IL17BR (interleukin 17 receptor beta) expression correlating with distant metastasis-free survival. For untreated patients with ER-positive, node-negative breast cancer, a HOXB13:IL17BR ratio of −2.0 results in an estimated risk of 15%, a HOXB13:IL17BR ratio of +2.0 in an estimated risk of 36% to develop recurrence at 5 years [65] and the cutoff point is set at 1.00 [66]. For the group of tamoxifen treated patients, a cut-point value of 0.06 for the HOXB13:IL17BR ratio is applied with this cutoff point having only predictive value for the treated ER-positive node-negative subgroup but not for the treated ER-positive node-positive subgroup [65, 67]. These results were supported by a large study on 1252 ER-positive primary breast tumors. The HOXB13:IL17BR ratio was found to be prognostic in lymph-node-negative untreated patients as had been reported in the initial study. The HOXB13:IL17BR expression ratio was assessed in tamoxifen-treated patients and showed significant association for progression free as well as post-relapse survival. Furthermore, in this study the HOXB13:IL17BR expression ratio showed a statistically significant association with disease outcome also in the group of lymph-node positive untreated patients [66], which, unfortunately, could not be validated in independent studies [67, 68].

To further delineate the role of the HOXB13:IL17BR expression ratio in the response towards tamoxifen therapy, the analysis of polymorphisms in the CYP2D6 gene encoding the enzyme that processes tamoxifen to its active metabolites endoxifen and 4-OH tamoxifen was combined with the HOXB13:IL17BR ratio into a CYP2D6:HOXB13/IL17BR risk factor. In ER-positive lymph-node negative breast cancer patients, women with 2 CYP2D6:HOXB13/IL17BR risk factors (decreased CYP2D6 metabolism and HOXB13/IL17BR high) had the shortest disease-free survival compared to those women with only one risk

factor or women with no CYP2D6:HOXB13/IL17BR risk factor at all (extensive CYP2D6 metabolizer and HOXB13/IL17BR low) [69]. As a stepwise increase of recurrence and death is observed in patients treated with tamoxifen if one or more CYP2D6:HOXB13/IL17BR risk factors are existent, the HOXB13:IL17BR ratio together with the CYP2D6 status might be used to allocate patients to endocrine therapy with or without concomitant chemotherapy.

### Prediction—response to cytotoxic therapy

Current approaches to determine the most effective therapeutic strategy for each individual patient are still limited [24, 70]. Over the last decade, therapy for cancer has gradually changed from applying unspecific cytotoxic treatment to a more target-oriented treatment that relies on specific biomarkers [70].

Attempts to predict treatment outcome according to different treatment options in breast cancer patients seems encouraging. Today, it is difficult to predict whether a certain chemotherapeutic approach will be effective for the individual patient. Although clinically very similar, two breast cancer patients might respond very differently to the same type of chemotherapy. In an attempt to solve this clinical problem, gene expression profiling (GEP) has been assessed in respect to prediction of response to chemotherapy in adjuvant and neoadjuvant settings.

Today, still the majority of patients with breast cancer are advised to receive chemotherapy when they are in an intermediary to high-risk situation for recurrence according to NIH and/or St. Gallen criteria. However, many patients are treated similarly, although it is already known that not all patients will respond due to a priori or de novo developed drug resistance. Neoadjuvant chemotherapy is perfectly suited for applying gene-expression profiling to develop signatures predicting in vivo responses towards chemotherapy, as outcomes in disease-free and overall survival are comparable between neoadjuvant and adjuvant chemotherapy [71]. Contemporaneous to the development of gene expression signatures for the prediction of prognosis in the untreated setting, signatures have been developed with the attempt to identify patients that may benefit from neoadjuvant chemotherapy. On the other hand, NPV for each regimen-specific genomic signature might identify patients who are unlikely to respond to a certain chemotherapy regimen, rendering them to more promising therapies.

Docetaxel is frequently used in the treatment of breast cancer and is one of the most active but also toxic and certainly not least expensive agents used to treat this disease. Although a significant proportion of breast cancer patients does not respond to docetaxel predictive factors

identified so far have not yet found the way into the clinical setting [72, 73]. Chang and collegues developed a gene-expression signature to predict benefit from taxane-based chemotherapy. 24 Patients with mainly locally advanced lymph-node positive breast cancer with mixed menopausal status that participated in a phase II study of neoadjuvant docetaxel were selected [74]. Sensitivity and resistance to neoadjuvant chemotherapy were defined as residual disease of less or more than 25% after neoadjuvant chemotherapy, respectively. A 92-gene predictive classifier with high expression of apoptosis-related and DNA damage-related genes was found in docetaxel-sensitive tumors. In leave-one-out cross-validation, this classifier showed a sensitivity of 90% to correctly classify effectiveness to docetaxel therapy. In a small validation set, the classifier correctly predicted response to chemotherapy in six of six patients, still the study cohort size is too small to obtain sufficient statistical power.

A diagnostic technique based on RT-PCR on breast cancer samples of 44 patients was used to generate another predictive gene signature for response to docetaxel chemotherapy [75]. In a second independent validation cohort, this group of 85 genes showed an overall accuracy of 80.7% with PPV and NPV of 73.3% and 90.9%, respectively. The 85-gene profile outcompeted clinicopathological parameters, none of which were significantly associated with tumor response to docetaxel.

In 45 patients with stage II or stage III breast cancer with only 14% reaching a complete pathologic response, a 22-gene signature consisting of angiogenesis-, proliferation- and invasion-related genes was found to be predictive for complete pathological remission in patients neoadjuvantly treated with 3 cycles of doxorubicin and six cycles of docetaxel [76]. Indeed the study showed the feasibility of RT-PCR based methods to explore candidate genes that correlate with pathologic complete response, but the 22-gene signature was not validated in a second independent study.

To predict at least partial pathological response to doxorubicin combined with cyclophosphamide, two gene expression signatures have been described [77, 78]. One classifier based on expression of only three genes (PRSS11, MTSS1 and CLPTM1) was established in a setting of four cycles neoadjuvant chemotherapy with doxorubicin and cyclophosphamide (AC) in patients with non-inflammatory, mostly advanced breast cancer [78]. This classifier was established in 44 samples (31 samples as training set and 13 samples in the validation set). In this small validation set of 13 samples, the 3-gene classifier only classified 11 patients correctly. The other classifier was based on the expression value of 253 genes involved in cell cycle, survival, stress response and ER-pathway. However, in leave-one-out cross validation this classifier classified only 67% of samples correctly into AC- sensitive or resistant cases [77].

The combination of docetaxel with trastuzumab in the treatment of HER2-positive breast cancer is common, but predictors of trastuzumab resistance are lacking. Based on microarray analysis of 25 tumors a 28-gene expression profile was identified and subsequently validated in 13 samples to predict complete pathologic response after docetaxel-trastuzumab treatment with an overall accuracy of 92%, with 100% sensitivity and 89% specificity [79]. Nonetheless, like the two aforementioned studies, also this study is significantly underpowered and the resulting classifier is likely to result from overfitting the data.

Gianni and colleagues investigated the role of the RS for prediction of response to chemotherapy in a neoadjuvant study of 89 patients with locally advanced breast cancer treated with doxorubicin and paclitaxel followed by 12 cycles of paclitaxel weekly, surgery and adjuvant cyclophosphamide, methotreaxate and 5-fluorouracil (CMF) for 4 cycles [80]. In this study, the RS was positively associated with the likelihood of pathologic complete response indicating patients with high RS likely benefitting from neoadjuvant chemotherapy. To assess whether the RS can indeed predict benefit from chemotherapy, 651 ER-positive lymph-node negative patients from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B20 trial were treated either with adjuvant tamoxifen or adjuvant tamoxifen plus cyclophosphamide, methotrexate, and fluorouracil (CMF) or methotrexate and fluorouracil (MF) [81]. Retrospectively, particularly high-risk patients with a RS of ≥31 benefitted from chemotherapy (distant recurrence free survival 60% with tamoxifen compared to 88% with tamoxifen plus chemotherapy) whereas low-risk patients basically did not. In 465 patients with locally advanced hormone receptor-positive breast cancer enrolled in the Eastern Cooperative Oncology Group (ECOG) E2197-trial treated with doxorubicin plus cyclophosphamide or docetaxel, plus tamoxifen if ER-positive, the RS was a highly significant predictor of recurrence in both lymph-node negative and lymph-node positive disease [82]. Moreover the RS predicted recurrence more accurately than standard clinicopathological features. 46% of patients were allocated by the RS to the low-risk group, resulting in a recurrence risk with chemotherapy of 3% if zero to one lymph-nodes and of 8% if two to three lymph-nodes were positive. This study did not include an arm without chemotherapy, but a study on postmenopausal women with locally advanced breast cancer found no benefit from adjuvant chemotherapy with cyclophosphamide, doxorubicin, and fluorouracil (CAF) when added to tamoxifen in patients with a low RS [61].

To evaluate the effect of chemotherapy on women with midrange risk of cancer recurrence the Eastern Cooperative Oncology Group (ECOG) coordinates the TAILORx (Trial Assigning IndividuaLized Options for Treatment) Trial

(Fig. 5). In this prospective study, patients are assigned to therapy according to their RS. Patients with a Recurrence Score of <11 (estimated 29% of the study population) receive antihormonal therapy, whereas patients with a Recurrence Score of >25 (estimated 27% of the study population) receive chemotherapy. About 44% of patients are considered to fall into the primary study group that contains patients with a midrange Recurrence score between 11 and 25. These patients, stratified into pre-, peri- or postmenopausal women, are randomized into a taxane-containing or a non-taxane containing chemotherapy to identify the best treatment option for this subset of patients.

Response to 4 cycles of T/FAC neoadjuvant chemotherapy might be predicted by a 74-gene profiler that was established in a small study on 42 patients receiving neoadjuvant paclitaxel and fluorouracil + doxorubicin + cyclophosphamide (T/FAC) chemotherapy [83]. The study cohort consisted of ER-receptor positive and ER-receptor negative patients with the majority of patients being HER2-negative. 48% of patients had no lymph-node involvement and 52% of patients had N1 or N2 disease. Among the 18 patients in the validation cohort the 74-gene profiler correctly predicted complete pathologic response in all three patients achieving complete remission thereby resulting in a positive predictive value of the positive 74-gene

profile of 100%. However only a small proportion of patients with a negative 74-gene profile were correctly diagnosed resulting in negative predictive value of only 73% and also this study is to small to draw a conclusion.

Another pharmacogenomic predictor for complete pathologic response to neoadjuvant T/FAC chemotherapy, the "DLDA-30" (Diagonal Linear Discriminant Analysis) gene profiler consisting of only 26 genes, was established based on the analysis of only 30 patients and was subsequently validated in a cohort of 51 patients [84]. The NPV of the DLDA-30 predictor was determined to be 96%. In addition, the DLDA-30 predictor correctly identified 92% of those patients achieving a pathological complete response, but many patients that are predicted by the DLDA-30 to have a cPR do not (PPV of 52%).

Taken together, all the aforementioned studies assessing response to chemotherapy have several important weaknesses: first, they are mainly based on much too small numbers of patients with too low statistical power considering the thousands of genes measured using array technology, resulting in overfitting of the data [85]. This results in very good predictive values in initial studies, yet as these results are mostly validated by cross-validation within the same dataset they are unlikely to be validated in larger validation studies. The second weakness is the lack of sufficiently large validation studies. In principle, none of
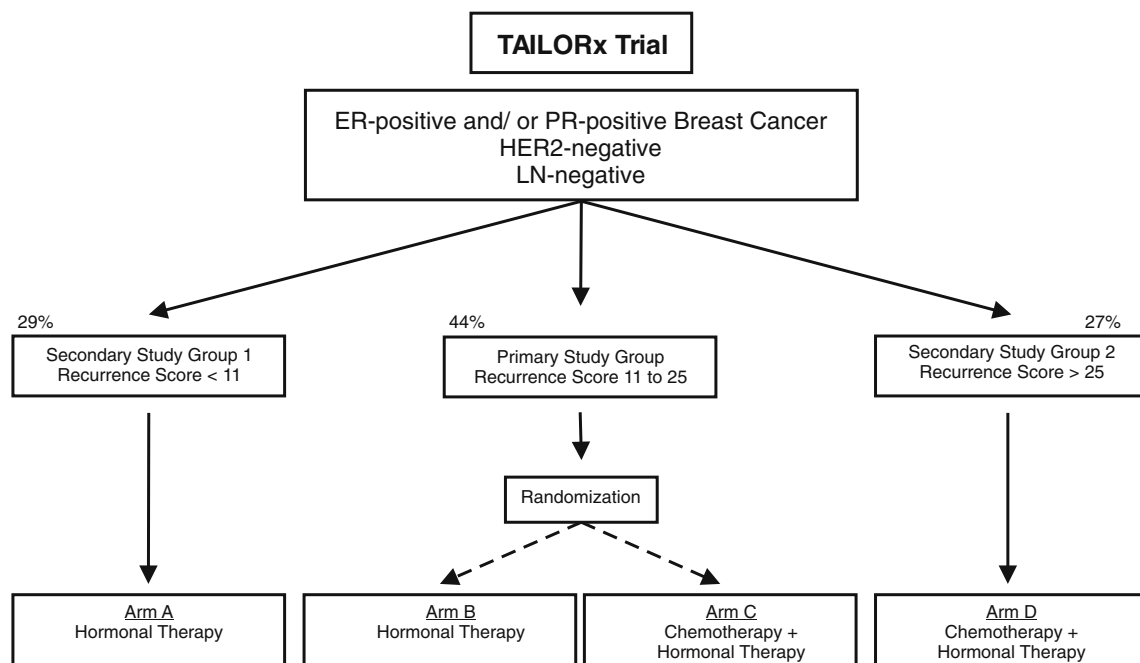


Fig. 5 Overview of the TAILORx Trial. The TAILORx trial has been designed to evaluate the role of intermediate RS in the assignment to adjuvant hormonal therapy alone in comparison to hormonal therapy in combination with chemotherapy. Patients with ER-positive LN-negative breast cancer are stratified according to their OncotypeDX Recurrence score. Patients with a RS of less than 11 are considered for hormonal therapy only. Patients with a RS of 25 or higher are assigned to chemotherapy plus hormonal therapy. Patients with a RS between 11 and 25 are randomly assigned to chemotherapy plus hormonal therapy (the standard treatment arm) or hormonal therapy alone (the experimental-treatment arm)

these approaches has been sufficiently validated, a prerequisite for clinical application. Another critical issue is the inclusion of both patients with ER-positive and ER-negative tumors in these studies, although it is well known that these subtypes respond differentially to neoadjuvant chemotherapy [86, 87].

In an attempt to translate gene expression signatures generated in vitro using chemotherapy sensitive cell lines into clinical application, single-agent drug sensitivity signatures were combined with FEC (fluorouracil, epirubicin and cyclophosphamide for six cycles) and T-ET (docetaxel for three cycles followed by epirubicin and docetaxel for three cycles) regimen specific signatures. Those were subsequently used to predict response in 212 patients with ER-negative breast cancer treated with an epirubicin based therapy in the EORTC 10994/BIG 00-01 trial. The FEC predictor was validated in a set of 66 ER-negative breast cancer patients (20% HER2 positive) treated with FEC from which 28 showed pathological complete responses. The FEC predictor predicted a pathological complete response correctly in 27 of 40 patients (PPV: 68%) and accurately identified 25 of 26 patients that did not respond to FEC chemotherapy (NPV: 96%). In parallel, a TET predictor was validated in 59 patients with ER-negative breast cancer (34% HER2 positive) treated with the taxane regimen docetaxel for three cycles followed by epirubicin and docetaxel (TET), among which 27 patients showed a pathological complete response. This predictor showed a PPV of 71% and a NPV of 92%. From this study, the authors concluded that selection of patients with expression of either the FEC or the TET predictor would allow reasonable allocation to the particular treatment. Allocation to FEC or TET based chemotherapy according to the respective predictor could increase proportion of patients with pathological complete response significantly from 44% to around 70% [88].

An attempt to increase the rate of pathological complete response (pCR) is the addition of gemcitabine and docetaxel to anthracyclin based chemotherapy, either combined (termed GEDoc) or dose dense and sequential (termed GEsDoc) with pCR rates in up to a quarter of patients [89, 90]. The use of gene expression profiling to predict benefit from this intensive chemotherapy was evaluated in 100 patients with ER-receptor positive and negative primary non-metastatic breast cancer [91]. Using gene expression data from patients with cPR in the GEDoc study as training set, a 512 gene expression profile that predicts a pCR after neoadjuvant systemic therapy containing gemcitabine, epirubicin and docetaxel was validated in patients undergoing GEDoc treatment. The established predictor showed a 88% overall accuracy with a high sensitivity of 78% and a specificity of 90%. However, even for this study secondary validation data are yet to be established.

Several studies have been reported describing the prediction of therapy response related to a certain chemotherapeutic regimen, e.g. the 74 gene profiler for response to T/FAC neoadjuvant chemotherapy, the 512 gene signature for gemcitabine/ docetaxel/ anthracyclin based chemotherapy and the FEC and TET predictor for epirubicin based chemotherapy combined with fluorouracil and cyclophosphamide or docetaxel, respectively [74, 83, 88, 91]. All of them have been identified as representative gene expression signatures that occur upon specific cytotoxic cell responses, and partially, it was demonstrated that the respective signature is specific for a certain combination of chemotherapeutic agents [88]. When designing a study to search for a gene signature that could predict response to neoadjuvant FEC treatment in patients with ER-negative tumors treated in the EORTC 10994/BIG 00-01 trial, Farmer and colleagues found a stroma-related gene signature, the "stromal metagene", whose expression was associated with significantly shorter relapse-free survival [37]. Interestingly, this signature could predict response not only to neoadjuvant FEC treatment, but also to T-FAC treatment (neoadjuvant chemotherapy with paclitaxel, 5-fluorouracil, doxorubicin and cyclophosphamide) from another independent cohort of ER-negative breast tumors. This suggests that the underlying biological response of the tumor microenvironment might be important for chemotherapy-dependent tumor eradication and the importance of the tumor stroma in metastasis, prognosis and response towards therapy is increasingly recognized [92–96]. As the "stromal metagene" showed its power in predicting treatment response, its prognostic power was assessed in an independent cohort of untreated patients from the Nederlands Kanker Instituut (NKI) and the Erasmus Medical Center (EMC) [37]. Whereas higher expression of the "stromal metagene" was associated with a significantly shorter relapse-free survival in patients treated with chemotherapy, it was unrelated to survival in the untreated patients. This underlines the predictive power of the "stromal metagene" expression signature rather than its prognostic role in untreated patients [37].

Since this gene expression signature seemed to be predictive for outcome to more than one chemotherapeutic regimen, the question arises whether the breast cancer intrinsic subtypes themselves respond differentially to chemotherapy. This applies a fortiori, as it is already known that tumor intrinsic factors influence chemotherapy efficiency. The expression of ER for example is negatively predictive for response to chemotherapy and the same might be suggested for its accompanying luminal A gene expression profile. Indeed, in 82 patients treated with T/FAC neoadjuvant chemotherapy basal-like and HER2[+] subgroups were associated with high rates of pCR, 45% and 45%, respectively. In contrast and in line with ER

receptor status, luminal tumors had a pathological CR rate of only 7% and no pCR was observed in the normal-like subclass [97]. The application of tumor intrinsic subtypes in predicting response to chemotherapy might be used in settings, where it is still difficult to find gene expression patterns that predict for therapy response [98, 99].

In conclusion, Patients that harbor a prognostic high-risk gene expression profile according to Fig. 3 and are assigned to undergo chemotherapy to minimize recurrence rate might be further allocated to tumor-specific chemotherapy. By extraction of gene expression profiles predictive for specific chemotherapies and merging to clinical patient parameters such as age, comorbidities, ECOG status, patients might be allocated to the treatment they profit most of (Fig. 4).

## Challenges

Enormous efforts have been undertaken and a high number of reports exist on prognostic tumor markers. However, the number of markers that are clinically useful is still small irrespective whether genomic or proteomic technologies are applied [100]. Several reasons account for this discrepancy, e.g. the lack of standardized technologies, study design with far too small patient size leading to overfitting thereby resulting in poor performance of established predictors in clinically meaningful validation studies, or even the lack of any meaningful validation studies [100]. Moreover, gene expression studies rely on the informative value of the whole specimen from which RNA is extracted and expression profiling is conducted. At this point, it is important to emphasize that tumors are a heterogeneous mixture of cells including the tumor cells with varying degree of differentiation but also inflammatory immune cells, surrounding stromal tissue and blood vessels. The amount of the respective cell type varies significantly not only between different tumor stages and grades, but also between tissues of different patients with tumors of the same histological subtype and grade. This clearly influences the designation to a particular prognostic group and consecutively the designation to one or the other therapy. It seems reasonable to dissect the tumor and isolate pure-tumor cell populations prior to gene expression profiling, but evidence is emerging that the interaction of tumor with the stroma and cells of the immune system plays a critical role in tumor progression and response towards chemotherapy [101, 102]. Furthermore, recent data emphasize the prognostic and predictive significance of stroma-related gene signatures. Finak et al. describe a SDPP (stroma-derived prognostic predictor), a 26-gene expression profile, which irrespective of standard clinical prognostic factors stratifies disease outcome [103] and predicts response to neoadjuvant chemotherapy with two different anthracycline-based regimens, FEC and T-FAC, in a large cohort of patients [37].

In addition to tumor intrinsic factors such as cellular composition and tumor differentiation, factors related to sample procurement also influence the overall gene expression profile. These include differences in sample preparation, selection of microarray platform and use of hybridization conditions. An important issue is the quality of RNA extracted from the tumor as well. RNA is unstable in tissue samples due to the high prevalence of RNAses, requiring quick freezing und processing of the sample. When validating the 70-gene prognosis signature in node-negative breast cancer patients, Bueno-de-Mesquita and colleagues reported up to 1/3 of samples to be excluded due to bad sample and/ or RNA quality [30]. In this regard, RT-PCR based tests as the OncotypeDX assay might be used in cases were only paraffin embedded tissue is available and if suited with comparable power regarding its prognostic and predictive capability. Unfortunately, no study to date has systemically compared the same patient-samples to multiple test assays to answer this important question. Pitfalls in RT-PCR based studies arise from other sources that might be as trivial as e.g the use of different primer pairs for detection. For the analysis of IL17BR expression, Ma et al. applied a primer set in the 3′ region revealed six time higher expression levels when directly comparing to a primer set in the 5′ region used by Reid et al. [65–67]. Nonetheless, levels of both IL17BR and HOXB13 correlated and consequently, the IL17BR:HOXB13 ratio was comparable [66].

When harmonizing studies on prognostic tumor markers, poor study design and analysis, assay variability and inadequate study reporting were identified as major barriers in the field of cancer diagnostics. This analysis lead to the development of the REMARK guidelines to encourage transparent and complete reporting on newly found prognostic markers [100]. The REMARK guidelines range from accurate description of patient characteristics to illustration of study design and statistical methods that have been used. Ideally, one should be able to compare the expression data obtained in any research facility at any time to any other data obtained in another facility at other time points using other microarray platforms.

## Perspective

Several approaches to deal with microarray data have been recently described, most of them being at the stage of translational research, but several ready to be implemented into clinical practice. In studies under the aegis of the TRANSBIG consortium, the predictive value of the 70-gene expression signature and the 76-gene expression

signature are as good as the best validated clinical tool Adjuvant!Online [29, 45]. Even more, due to their higher specificity, they both seem to identify low-risk patients better implying a potential to reduce unnecessary chemotherapy [104]. In this regard, it is important to note that the OncotypeDX recurrence score and the Adjuvant! Online tool estimate different parameters. Whereas both the Mammaprint signature and the OncotypeDX estimate only the risk for distant recurrence (risk of distant metastasis), the Adjuvant! Online tool estimates risk for all causes of recurrence (local, regional, distant recurrence and contralateral breast cancer), making an exact comparison between Mammaprint, OncotypeDX and Adjuvant! Online difficult, if not impossible.

Nonetheless, the majority of signatures are developed from distinct sets of mainly small patient populations with limited validation and follow-up, as obvious in the studies that have established the 92-gene and the 85-gene classifier predictive for docetaxel treatment, the 22-gene classifier predictive for doxorubicin-docetaxel treatment, the 3-gene classifier predictive for doxorubicin-cyclophosphamide treatment, the 28-gene classifier predictive for docetaxel-trastuzumab treatment, the 74-gene and the "DLDA-30" profile classifier predictive for T/FAC chemotherapy, the FEC- and TET- predictors, the GEDoc predictor and finally both the stromal metagene and the SDPP.

This holds partially true for the well established 70-gene profiler Mammaprint, a signature superior to traditional clinical predictors within the follow-up of the initial study it was derived from. However, when applying this 70-gene profile to longer follow-up, the gene-expression signature show heterogeneous behavior, indicating that different mechanisms might be responsible for early (within 5 years) and late (beyond 5 years) distant metastasis [104].

Interestingly, and discordant with the fear of many researchers, it was shown that different chemotherapeutic agents can elicit similar response signatures. Although this naturally has to be validated on larger cohorts, it reminds one that we, even with "undirected" chemotherapy such as anthracyclins, seem to target specific pathways rather than evenly administering cytotoxicity like with a watering can.

Other important conclusions can be drawn from the study by Bonnefoi et al.. First, under certain circumstances, gene signatures calculated from cell line data can be recovered in the clinical "in vivo setting" and second, certain gene expression data obtained from different gene expression platforms might be integrated after biostatistical corrections are performed [88].

Besides developing classifiers for prognosis and prediction, GEP also can be used to identify genes that e.g. mediate resistance to specific chemotherapeutic agents [105, 106]. Moreover, GEP might guide the pathologist, when correctly allocating histologically graded intermediate grade 2 breast carcinoma into genomic grade 1 or 3 [52]. GEP has the potential to substantially refine cancer prognosis well beyond what is currently possible with the clinicopathologic indicators. In part, GEP is ripe for introduction into the clinic and may guide systemic therapy in the future. This is especially true for the well validated 70-gene prognosis score (MammaPrint) and the Oncotype DX recurrence score, whereas caution has to be applied to the many genomics signatures assessed in only small subsets of patients likely not reflecting the whole spectrum of disease and therefore not being representative [107]. Moreover, beside allocation to the main 5 subtypes of breast cancer, already today GEP can be used to identify the molecular basis of the disease. For example, Hedenfalk et al. used GEP to study seven spontaneous and 15 hereditary breast adenocarcinomas with mutations in either BRCA1 or BRCA2. They were able to identify a number of differentially expressed genes between BRCA1-mutated and BRCA2-mutated tumors and used these genes to accurately identify breast cancer samples that harbored these genetic mutations [108].

The information deciphered by GEP methods might not only accelerate identifying novel molecular targets, but might—by providing the clinician a description of the tumors pathology and chemosensitivity—accompany the patient through her exertive walk against the cancer.

## References

1. Wolff AC, Davidson NE. Primary systemic therapy in operable breast cancer. J Clin Oncol. 2000;18:1558–69.
2. Chia S, Swain SM, Byrd DR, Mankoff DA. Locally advanced and inflammatory breast cancer. J Clin Oncol. 2008;26:786–90.
3. Golubnitschaja O. Predictive diagnostics and personalized treatment: dream or reality. New York: Nova Biomedical Books; 2009.
4. Smith IE, Dowsett M. Aromatase inhibitors in breast cancer. N Engl J Med. 2003;348:2431–42.
5. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Herceptin Adjuvant (HERA) Trial Study Team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005;353:1659–72.
6. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer Jr CE, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med. 2005;353:1673–84.
7. Hudis CA. Trastuzumab—mechanism of action and use in clinical practice. N Engl J Med. 2007;357:39–51.
8. Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. N Engl J Med. 2006;355:2733–43.
9. Blackwell KL, Burstein HJ, Storniolo AM, Rugo H, Sledge G, Koehler M, et al. Randomized study of lapatinib alone or in

combination with trastuzumab in women with ErbB2-positive, trastuzumab-refractory metastatic breast cancer. J Clin Oncol. 2010;28:1124–30.

10. Di Leo A, Gomez HL, Aziz Z, Zvirbule Z, Bines J, Arbushites MC, et al. Phase III, double-blind, randomized study comparing lapatinib plus paclitaxel with placebo plus paclitaxel as first-line treatment for metastatic breast cancer. J Clin Oncol. 2008;26:5544–52.

11. Brandes AA, Franceschi E, Tosoni A, Degli Esposti R. Trastuzumab and lapatinib beyond trastuzumab progression for metastatic breast cancer: strategies and pitfalls. Expert Rev Anticancer Ther. 2010;10:179–84.

12. Braun M, Fountoulakis M, Papadopoulou A, Vougas K, Seidel I, Höller T, et al. Down-regulation of microfilamental network-associated proteins in leukocytes of breast cancer patients: potential application to predictive diagnosis. Cancer Genomics Proteomics. 2009;6:31–40.

13. Li W. How many genes are needed for early detection of breast cancer, based on gene expression patterns in peripheral blood cells? Breast Cancer Res. 2005;7:E5.

14. Sharma P, Sahni NS, Tibshirani R, Skaane P, Urdal P, Berghagen H, et al. Early detection of breast cancer based on gene-expression patterns in peripheral blood cells. Breast Cancer Res. 2005;7:R634–44.

15. Yeghiazaryan K, Mamlouk S, Trog D, Moenkemann H, Braun M, Kuhn W, et al. Irradiated breast cancer patients demonstrate subgroup-specific regularities in protein expression patterns of circulating leukocytes. Cancer Genomics Proteomics. 2007;4:411–8.

16. Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, Simpson JT, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature 2009;462:1005–10.

17. Zhao Q, Caballero OL, Levy S, Stevenson BJ, Iseli C, de Souza SJ, et al. Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. Proc Natl Acad Sci USA. 2009;106:1886–91.

18. Cheang MC, van de Rijn M, Nielsen TO. Gene expression profiling of breast cancer. Annu Rev Pathol. 2008;3:67–97.

19. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.

20. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001;98:10869–74.

21. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci USA. 2003;100:8418–23.

22. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA. 2003;100:10393–8.

23. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci USA. 2003;100:15901–5.

24. 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;220:263–80.

25. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.

26. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002;347:1999–2009.

27. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Senn HJ. 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.

28. Eifel P, Axelson JA, Costa J, Crowley J, Curran Jr WJ, Deshler A, et al. National institutes of health consensus development conference statement: adjuvant therapy for breast cancer, November 1–3, 2000. J Natl Cancer Inst. 2001;93:979–89.

29. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, 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.

30. Bueno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. Breast Cancer Res Treat. 2009;117:483–95.

31. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1–3 positive lymph nodes in an independent validation study. Breast Cancer Res Treat. 2009;116:295–302.

32. Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, et al. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol. 2008;26:3324–30.

33. Mook S, Schmidt MK, Weigelt B, Kreike B, Eekhout I, van de Vijver MJ, et al. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. Ann Oncol. 2010;21:717–22.

34. Wittner BS, Sgroi DC, Ryan PD, Bruinsma TJ, Glas AM, Male A, et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. Clin Cancer Res. 2008;14:2988–93.

35. Hayes DF. Follow-up of patients with early breast cancer. N Engl J Med. 2007;356:2505–13.

36. Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, Bakx N, et al. Converting a breast cancer microarray signature into a high-throughput diagnostic test. BMC Genomics. 2006;7:278.

37. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009;15:68–74.

38. Bueno-de-Mesquita JM, van Harten WH, Retel VP, van't Veer LJ, van Dam FS, Karsenberg K, et al. 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;8:1079–87.

39. Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, et al. Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. Nat Clin Pract Oncol. 2006;3:540–51.

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

41. Esteva FJ, Sahin AA, Cristofanilli M, Coombes K, Lee SJ, Baker J, et al. Prognostic role of a multigene reverse transcriptase-PCR assay in patients with node-negative breast cancer not receiving adjuvant systemic therapy. Clin Cancer Res. 2005;11:3315–9.

42. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, 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.

43. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365:671–9.

44. Foekens JA, Atkins D, Zhang Y, Sweep FC, Harbeck N, Paradiso A, 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.

45. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANS-BIG multicenter independent validation series. Clin Cancer Res. 2007;13:3207–14.

46. Eichhorn PJ, Baselga J. HER2 signatures in breast cancer: ready to go to print? J Clin Oncol. 2010;28:1809–10.

47. Staaf J, Ringnér M, Vallon-Christersson J, Jönsson G, Bendahl PO, Holm K, et al. Identification of subtypes in human epidermal growth factor receptor 2-positive breast cancer reveals a gene signature prognostic of outcome. J Clin Oncol. 2010;28:1813–20.

48. Alexe G, Dalgin GS, Scanfeld D, Tamayo P, Mesirov JP, DeLisi C, et al. High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates. Cancer Res. 2007;67:10669–76.

49. Ramaswamy S, Perou CM. DNA microarrays in breast cancer: the promise of personalised medicine. Lancet 2003;361:1576–7.

50. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci USA. 2005;102:3738–43.

51. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol. 2004;2:E7.

52. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst. 2006;98:262–72.

53. Liu R, Wang X, Chen GY, Dalerba P, Gurney A, Hoey T, et al. The prognostic role of a gene signature fromtumorigenic breast-cancer cells. N Engl J Med. 2007;356:217–26.

54. Group, EBCTC. 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.

55. Fisher B, Jeong JH, Bryant J, Anderson S, Dignam J, Fisher ER. National Surgical Adjuvant Breast and Bowel Project randomised clinical trials. Treatment of lymph-node-negative, oestrogen-receptor-positive breast cancer: long-term findings from National Surgical Adjuvant Breast and Bowel Project randomised clinical trials. Lancet 2004;364:858–68.

56. Bardou VJ, Arpino G, Elledge RM, Osborne CK, Clark GM. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. J Clin Oncol. 2003;21:1973–9.

57. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell. 2004;5:607–16.

58. Clarke R, Liu MC, Bouker KB, Gu Z, Lee RY, Zhu Y, et al. Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. Oncogene 2003;22:7316–39.

59. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. J Clin Oncol. 2008;26:721–8.

60. Albain KS, Barlow WE, Ravdin PM, Farrar WB, Burton GV, Ketchel SJ, et al. Breast Cancer Intergroup of North America.

Adjuvant chemotherapy and timing of tamoxifen in postmeno-pausal patients with endocrine-responsive, node-positive breast cancer: a phase 3, open-label, randomised controlled trial. Lancet 2009;374:2055–63.

61. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Breast Cancer Intergroup of North America. 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.

62. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. 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;28:1829–34.

63. Andre F, Delaloge S. First-generation genomic tests for breast cancer treatment. Lancet Oncol. 2010;11:6–7.

64. Lo SS, Mumby PB, Norton J, Rychlik K, Smerage J, Kash J, et al. Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol. 2010;28:1671–6.

65. Ma X-J, Hilsenbeck SG, Wang W, Ding L, Sgroi DC, Bender RA, et al. The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. J Clin Oncol. 2006;24:4611–9.

66. Jansen MP, Sieuwerts AM, Look MP, Ritstier K, Meijer-van Gelder ME, van Staveren IL, 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.

67. Reid JF, Lusa L, De Cecco L, Coradini D, Veneroni S, Daidone MG, et al. Limits of predictive models using microarray data for breast cancer clinical treatment outcome. J Natl Cancer Inst. 2005;97:927–30.

68. Goetz MP, Suman VJ, Ingle JN, Nibbe AM, Visscher DW, Reynolds CA, 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.

69. Goetz MP, Suman VJ, Couch FJ, Ames MM, Rae JM, Erlander MG, 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.

70. van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. Nature 2008;452:564–70.

71. Bear HD, Anderson S, Smith RE, Geyer Jr CE, Mamounas EP, Fisher B, et al. Sequential preoperative or postoperative docetaxel added to preoperative doxorubicin plus cyclophospha-mide for operable breast cancer: National Surgical Adjuvant Breast and Bowel Project Protocol B-27. J Clin Oncol. 2006;24:2019–27.

72. Egawa C, Miyoshi Y, Takamura Y, Taguchi T, Tamaki Y, Noguchi S. Decreased expression of BRCA2 mRNA predicts favorable response to docetaxel in breast cancer. Int J Cancer. 2001;95:255–9.

73. Hasegawa S, Miyoshi Y, Egawa C, Ishitobi M, Tamaki Y, Monden M, et al. Mutational analysis of the class I beta-tubulin gene in human breast cancer. Int J Cancer. 2002;101:46–51.

74. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003;362:362–9.

75. Iwao-Koizumi K, Matoba R, Ueno N, Kim SJ, Ando A, Miyoshi Y, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol. 2005;23:422–31.

76. Mina L, Soule SE, Badve S, Baehner FL, Baker J, Cronin M, et al. Predicting response to primary chemotherapy: gene expression profiling of paraffin-embedded core biopsy tissue. Breast Cancer Res Treat. 2007;103:197–208.

77. Cleator S, Tsimelzon A, Ashworth A, Dowsett M, Dexter T, Powles T, et al. Gene expression patterns for doxorubicin (Adriamycin) and cyclophosphamide (Cytoxan) (AC) response and resistance. Breast Cancer Res Treat. 2006;95:229–33.

78. Folgueira MA, Carraro DM, Brentani H, Patrão DF, Barbosa EM, Netto MM, et al. Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. Clin Cancer Res. 2005;11:7434–43.

79. Vegran F, Boidot R, Coudert B, Fumoleau P, Arnould L, Garnier J, et al. Gene expression profile and response to trastuzumab-docetaxel-based treatment in breast carcinoma. Br J Cancer. 2009;101:1357–64.

80. Gianni L, Zambetti M, Clark K, Baker J, Cronin M, Wu J, 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.

81. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, 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.

82. Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley 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.

83. Ayers M, Symmans WF, Stec J, Damokosh AI, Clark E, Hess K, 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.

84. Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol. 2006;24:4236–44.

85. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. Proc Natl Acad Sci USA. 2006;103:5923–8.

86. Guarneri V, Broglio K, Kau SW, Cristofanilli M, Buzdar AU, Valero V, et al. Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors. J Clin Oncol. 2006;24:1037–44.

87. Colleoni M, Viale G, Zahrieh D, Pruneri G, Gentilini O, Veronesi P, et al. Chemotherapy is more effective in patients with breast cancer not expressing steroid hormone receptors: a study of preoperative treatment. Clin Cancer Res. 2004;10:6622–8.

88. Bonnefoi H, Potti A, Delorenzi M, Mauriac L, Campone M, Tubiana-Hulin 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.

89. Schneeweiss A, Schuetz F, Rudlowski C, Hahn M, Lauschner I, Sinn HP, et al. Dose-dense primary systemic chemotherapy with gemcitabine plus epirubicin sequentially followed by docetaxel for early breast cancer: final results of a phase I/II trial. Anticancer Drugs. 2005;16:1023–8.

90. Schneeweiss A, Huober J, Sinn HP, von Fournier D, Rudlowski C, Beldermann F, et al. Gemcitabine, epirubicin and docetaxel as primary systemic therapy in patients with early breast cancer: results of a multicentre phase I/II study. Eur J Cancer. 2004;40:2432–8.

91. Thuerigen O, Schneeweiss A, Toedt G, Warnat P, Hahn M, Kramer H, et al. Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer. J Clin Oncol. 2006;24:1839–45.

92. Zitvogel L, Apetoh L, Ghiringhelli F, Kroemer G. Immunological aspects of cancer chemotherapy. Nat Rev Immunol. 2008;8:59–73.

93. Bergamaschi A, Tagliabue E, Sørlie T, Naume B, Triulzi T, Orlandi R, et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. J Pathol. 2008;214:357–67.

94. Wennmalm K, Ostman A, Bergh J. Stromal signature identifies basal breast cancers. Nat Med. 2009;15:237–8.

95. Patocs A, Zhang L, Xu Y, Weber F, Caldes T, Mutter GL, et al. Breast-cancer stromal cells with TP53 mutations and nodal metastases. N Engl J Med. 2007;357:2543–51.

96. McMillin DW, Delmore J, Weisberg E, Negri JM, Geer DC, Klippel S, et al. Tumor cell-specific bioluminescence platform to identify stroma-induced changes to anticancer drug activity. Nat Med. 2010;16:483–9.

97. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res. 2005;11:5678–85.

98. Hannemann J, Oosterkamp HM, Bosch CA, Velds A, Wessels LF, Loo C, et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol. 2005;23:3331–42.

99. Sørlie T, Perou CM, Fan C, Geisler S, Aas T, Nobel A, et al. Gene expression profiles do not consistently predict the clinical treatment response in locally advanced breast cancer. Mol Cancer Ther. 2006;5:2914–8.

100. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M. Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies. J Clin Oncol. 2005;23:9067–72.

101. Bhowmick NA, Moses HL. Tumor-stroma interactions. Curr Opin Genet Dev. 2005;15:97–101.

102. Kim JB, Stein R, O'Hare MJ. Tumour-stromal interactions in breast cancer: the role of stroma in tumourigenesis. Tumor Biol. 2005;26:173–85.

103. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. Nat Med. 2008;14:518–27.

104. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat Rev Cancer. 2007;7:545–53.

105. Honma K, Iwao-Koizumi K, Takeshita F, Yamamoto Y, Yoshida T, Nishio K, et al. RPN2 gene confers docetaxel resistance in breast cancer. Nat Med. 2008;14:939–48.

106. Rhodes DR, Ateeq B, Cao Q, Tomlins SA, Mehra R, Laxman B, et al. AGTR1 overexpression defines a subset of breast cancer and confers sensitivity to losartan, an AGTR1 antagonist. Proc Natl Acad Sci USA. 2009;106:10284–9.

107. Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer. 2004;4:309–14.

108. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, et al. Gene-expression profiles in hereditary breast cancer. N Engl J Med. 2001;344:539–48.