EUROPEAN JOURNAL OF CANCER 45 (2009) 1733–1743

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Current Perspective

Predictive signatures for chemotherapy sensitivity in breast cancer: Are they ready for use in the clinic?

Hervé Bonnefoi[a,*], *Craig Underhill*[a], *Richard Iggo*[a], *David Cameron*[b]

[a]*Bergonié Cancer Institute and University of Bordeaux, 229 Cours de l'Argonne, 33076 Bordeaux Cedex, France*
[b]*NCRN Co-ordinating Centre, University of Leeds, Leeds, UK*

ARTICLE INFO

Article history:
Received 9 April 2009
Received in revised form 21 April 2009
Accepted 30 April 2009
Available online 26 May 2009

Keywords:
Breast cancer
Adjuvant
Chemotherapy
Gene signature
Predictive
Biomarkers
Chemosensitivity

ABSTRACT

Markers that predict the sensitivity of tumours to chemotherapy must address two questions: (a) which tumours are more likely to respond to chemotherapy? and (b) what is the optimal chemotherapy regimen for a specific tumour or group of tumours? To answer these questions will require markers of general chemosensitivity and drug-specific chemosensitivity, respectively. Beyond these fundamental questions lies an important practical question: are the predictive markers in the current literature ready for routine clinical use? The focus of this paper is to address this practical question. We will first review retrospective trials that have reported promising chemotherapy signatures, presenting in a comprehensive manner for the non bio-informatician the different methods used so far. In addition, we will summarise prospective trials (either ongoing or under development) designed to test the multigene classifiers currently thought to predict chemosensitivity. Finally, we will discuss why microarray studies have so far failed to identify new targets, and how we might be able to improve on these results through large-scale genotyping of tumours.

1. Introduction

Identification of markers that predict chemosensitivity is a research priority. The aim is to answer two very different questions: (a) can we use gene signatures to identify tumours which are more likely to respond to chemotherapy? and (b) when chemotherapy is indicated, what is the optimal chemotherapy regimen for a specific tumour or group of tumours? The predictive markers which answer these questions are likely to be different; for the sake of simplicity we will describe them in this review in two categories: markers predicting general chemosensitivity (meaning that a tumour is sensitive to any chemotherapy or to a wide range of chemotherapeutic drugs) and markers predicting drug-specific chemosensitivity (meaning that a tumour is sensitive to a specific class of agents). Of note, the identification of 'molecular features that indicate the optimal chemotherapy regimen' was considered a top priority in a recent internet-based consultation of 420 breast cancer researchers (clinicians, scientists, academics and pathologists).[1]

To date, predictive markers have been analysed either as single markers (for example, proliferation markers, hormone receptors, *HER2* and *p53*) or in groups, commonly referred to as gene signatures, metagenes, multigene biomarkers, multigene predictors or multigene classifiers. If we take oestrogen receptor (ER) status as an example of a single marker, several

* *Corresponding author:* Tel.: +33 5 56 33 78 17; fax: +33 5 56 33 33 30.
E-mail addresses: bonnefoi@bergonie.org, Quincy@bergonie.org (H. Bonnefoi).
0959-8049/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2009.04.036

studies have published provocative, but often contradictory, results regarding the value of ER status in predicting the benefit of adding taxanes to or after an anthracycline-based chemotherapy regimen.[2–6] Likely explanations for these apparently contradictory results are the heterogeneity of the ER-positive tumours included in these trials and the use of different cut-offs to define ER status. Multigene classifiers are more robust than single gene classifiers because random variation may negate the predictive information in some of the samples tested with only a single marker. Thus we have chosen in this review to concentrate on the literature assessing the predictive value of multigene classifiers. Two recent publications have comprehensively reviewed gene expression signatures in breast cancer but none has focused on the complex issue of their potential role in predicting chemotherapy sensitivity.[7,8]

Firstly, we will review retrospective trials that have identified promising multigene classifiers of chemotherapy sensitivity. Secondly, we will summarise prospective trials (either ongoing or under development) aiming to test the ability of multigene classifiers to predict chemosensitivity. Simon emphasised that 'a multigene biomarker can be a function that provides a continuous risk score rather than a class identifier'.[9] This difference is particularly important when aiming to prospectively validate a marker since the cut-off thresholds defining different classes (for example, good and bad responders) must be chosen in advance. As suggested by Simon, we prefer the phrase 'multigene classifier' rather than 'multigene biomarker' particularly when discussing prospective trials. Thirdly, we will discuss why, to date, gene expression signatures have failed to identify new targets and how we might be able to improve on these results.

2. Retrospective trials

2.1. *Predictive multigene classifiers of general chemosensitivity*

2.1.1. *The '21-gene recurrence score' (Oncotype DX™)*

Paik and collaborators developed a 21-gene recurrence score (RS) as a prognostic tool for predicting 10-year survival in a population of patients with early breast cancer.[10] Multiple quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) assays were used to quantify gene expression in formalin-fixed paraffin-embedded tissue. Initially, the correlation of gene expression with the likelihood of distant recurrence was studied in a series of 447 patients with node-negative, hormone receptor-positive breast cancer treated with tamoxifen in the National Surgical Adjuvant Breast and Bowel Project (NSABP) trial B-20. A panel of 16 cancer-related genes and 5 reference genes was selected and used to create a 21-gene recurrence score which increases smoothly from low, through intermediate, to high values. The recurrence score was then tested (some would say validated) in a series of 668 node-negative patients treated with tamoxifen alone in NSABP trial B-14. The rate of distant recurrence was 6.8% in the low recurrence score group, 14.3% in the intermediate score group and 30.5% in the high recurrence score group ($p < 0.001$).

Subsequently, two retrospective studies have reported that this same 21-gene recurrence score has predictive value for chemosensitivity.[11,12] The predictive value of the 21-gene recurrence score was assessed in 651 patients with node-negative, hormone receptor-positive tumours in NSABP trial B-20 randomised to tamoxifen alone ($n = 227$) or tamoxifen plus chemotherapy (methotrexate–fluorouracil or methotrexate–fluorouracil–cyclophosphamide) ($n = 424$).[11] A high recurrence score predicted benefit from chemotherapy (hazard ratio (HR) = 0.26; 95% confidence interval (CI) = 0.13–0.53), with little or no benefit from chemotherapy in the low and intermediate recurrence score groups.

The predictive value of the 21-gene recurrence score was also assessed in a subset of patients more than 50 years old with node-positive hormone receptor-positive tumours included in the SWOG 8814 trial.[12] In the SWOG 8814 trial, patients were randomised to receive either tamoxifen alone ($n = 361$); fluorouracil, doxorubicin and cyclophosphamide (CAF) followed by tamoxifen for 5 years ($n = 566$); or concurrent CAF and tamoxifen ($n = 550$). The 21-gene recurrence score was assessed in 367 of these patients. The score was low (<18) in 40%, intermediate (18–30) in 28% and high (>31) in 32% of patients. The addition of chemotherapy to tamoxifen resulted in no difference in disease-free survival (DFS) or overall survival (OS) in the low recurrence score group, but a clear benefit in DFS and OS in the high recurrence score group. There appeared to be a benefit as well for patients in the intermediate recurrence score group, but the confidence intervals were wide because of the small sample size. The results of this study, not yet fully published, suggest that the 21-gene recurrence score can identify one group of patients with node-positive, hormone receptor-positive disease who may derive no benefit from chemotherapy and another group who derive a larger benefit from chemotherapy than previously thought. If confirmed in further series of node-positive patients, this would seriously challenge the contemporary paradigm that all node-positive patients should receive chemotherapy.

2.1.2. *The '70-gene signature' (Mammaprint™)*

The '70-gene signature' developed by the Amsterdam group stratifies patients into poor prognosis and good prognosis groups.[13] Its prognostic value has been validated in cohorts of patients with node-negative[14,15] and node-positive tumours.[14,16] In a pooled multivariate analysis of two series of patients with node-positive tumours, the prognostic value of the gene signature was confirmed in a multivariate analysis (HR 5.50, 95% CI 1.47–20.62, $p = 0.01$).[16]

In addition, one retrospective study has suggested that the 70-gene signature might also predict the response to neoadjuvant chemotherapy. The signature was assessed in a series of 167 patients with tumours greater than 5 cm or clinically positive nodes.[17] Pathological complete response (pCR) after neoadjuvant chemotherapy was used as a surrogate for chemosensitivity. None of the patients with a good signature ($n = 144$) achieved a pCR (0/23), whereas 20% (29/144) of patients with a bad signature had a pCR. Consequently, the authors concluded that patients with a good signature would be unlikely to respond to chemotherapy.

Table 1 – Summary of trials aiming to identify predictive signatures of specific chemosensitivity.

	N total (learning set/validation set)	Chemotherapy treatment	Surrogate for chemosensitivity	Molecular biology method	Bioinformatics method	Signature identified	NPV[(a)]	PPV	Accuracy
Chang et al. (2003)	30 (24/6)	Docetaxel	Clinical response	cDNA chip	Supervised analysis	Yes (92 probesets)	83%	92%	88%
Ayers et al. (2004)	42 (24/18)	P ->FAC	pCR	cDNA chip	Supervised analysis	Yes (74 probesets)	73%	100% (3/3)	78%
Iwao-Koizumi et al. (2005)	70 (44/26)	Docetaxel	Clinical response	ATAC-PCR	Supervised analysis	Yes (85 genes)	90.9%	73.3%	80.7%
Hanneman et al. (2005)	46 (46/–)	AC or AD	pCR	cDNA chip	Supervised analysis	No	–	–	–
Gianni et al. (2005)	171(89/82)	AP->P	pCR	Q RT-PCR(1) and cDNA chip(2)	Supervised analysis	Yes (86 genes)	–	–	–
Hess et al. (2006)	133 (82/51)	P->FAC	pCR	cDNA chip	Supervised analysis	Yes (30 probesets)	96%	52%	76%
Thuerigen et al. (2006)	100 (52/48)	GE->Doc or GEDoc	pCR	cDNA chip	Supervised analysis	Yes (512 probesets)	95%	64%	78%
Cleator et al. (2006)	40 (40/0)	AC	Clinical response	cDNA chip	–	Yes	–	–	–
Potti et al. (2006)	51 (–/51)	T->FAC	pCR		*In vitro* signatures	Yes	94%	61.1%	82.3%
Bonnefoi et al. (2007)	125 (–/125)		pCR	cDNA chip	*In vitro* signatures	Yes			
	66	FEC					96%	68%	79%
	59	T->ET					92%	71%	80%
Chang et al. (2008)	72 (72/–)	Docetaxel	Clinical response	Q RT-PCR	–	Yes	–	–	–
Farmer et al. (2009)	63 (63/–)	FEC	pCR	CDNA chip	Metagenes analysis	Yes	81%	57%	65%

P->FAC: weekly paclitaxel × 12 followed by 5-fluorouracil, adriamycin, cyclophosphamide (FAC) every 3 weeks; AC: adriamycin, cyclophosphamide × 6; AD: doxorubicin docetaxel × 6; AP->P: paclitaxel and doxorubicin × 3 followed by weekly paclitaxel × 12; GE->Doc: gemcitabine, epirubicin × 5 followed by docetaxel × 4; GEDoc: gemcitabine, epirubicin, docetaxel × 6; FEC: 5-fluorouracil, epirubicin, cyclophosphamide × 6; T->ET: docetaxel (every 3 weeks) × 3 followed by epirubicin and docetaxel every 3 weeks × 3; pCR: pathological complete response; cDNA: complementary DNA; ATAC-PCR: adaptor-tagged competitive polymerase chain reaction; Q RT-PCR: quantitative reverse-transcriptase polymerase chain reaction method; (1) used in 89 patients as a 1st step; (2) used in 82 patients as a 2nd step; NPV: negative predictive value; and PPV: positive predictive value.

(a) NPV, PPV and accuracy are calculated on the validation set.

2.2. *Predictive multigene classifiers of specific chemosensitivity*

2.2.1. *Gene expression studies*

Several studies have attempted to identify multigene classifiers of response to specific chemotherapeutic agents or regimens[18–29] (Table 1). The common objective of these studies was to predict the clinical or pathological response to neoadjuvant chemotherapy by analysing fine needle aspiration or core biopsies taken before initiation of chemotherapy. The patients were all treated in phase 2 trials with the exception of the EORTC study, which analysed a subset of patients enrolled in a phase 3 trial comparing two chemotherapy regimens.[27] All of the studies used gene expression microarrays, except for three trials that used qRT-PCR[20,22,28] (Table 1).

2.2.1.1. *Techniques for construction of multigene classifiers.*

Conventional supervised analysis

Conventional supervised analysis compares the expression profiles of responsive tumours with the profiles of non-responsive tumours in order to identify a list of discriminant genes. This is a classic statistical technique which has been used in many other situations, for example, to identify prognostic signatures.[13,30] The population is divided into a training set and a validation set. The signature is developed in the training set and its performance is then tested in the validation set. In the training set, the most discriminant genes are commonly identified by some form of *t*-test comparing responders with non-responders. The size of the training and validation sets need not be equal: in the extreme case, the validation set contains only a single sample (so-called 'leave-one-out cross-validation'). In this case, one sample is left out, the analysis is run on the remaining samples to create the signature, and the status of the omitted sample is then predicted. The most promising signature from the learning set is used in the validation set to assess the performance (normally expressed as the sensitivity, specificity, negative predictive value, positive predictive value and accuracy). The MD Anderson Cancer Center (MDACC) group has extensive experience with this method.[19,23] In a recent publication they reported results comparing 20 different algorithms to select the discriminant genes, optimise the number of genes in the signature, and weight them to create the predictor.[23] In their experience, the best results were obtained with 30 probesets using diagonal linear discriminant analysis (DLDA). The main weakness of this method is 'overfitting', a term used to describe a signature that works perfectly in the training set, but has little predictive power in an independent validation set. This happens when the genes picked for the signature reflect characteristics of the tumours in the training set that turn out not to be relevant to the response. Indeed, modern algorithms can obtain perfect response predictions by overfitting even from random data.[31] A further problem is that cross-validation, such as the leave-one-out cross-validation procedure described above, does not validate any particular gene signature. Instead, it tests the algorithm used to produce the signature. The gene signature tested is normally different in every cycle of cross-validation. It only makes sense to claim that cross-validation tests a particular signature if it can be demonstrated that the same genes were selected in every cycle of cross-validation. The best way to avoid overfitting is to test signatures in a completely independent validation set, for example, patients treated in another hospital or study.

In vitro signature analysis

There is no *a priori* reason to use clinical samples in the training set. A group at Duke University has pioneered the use of hybrid techniques that combine microarray data from cell lines and tumours into a single matrix.[26] The Duke approach requires the use of sophisticated normalisation techniques to create a single combined dataset and powerful algorithms to project *in vitro* drug sensitivity data from the cell lines onto the tumours. They first selected cell lines resistant or sensitive to a variety of cytostatic drugs and identified genes associated with *in vitro* drug sensitivity by Bayesian binary regression analysis. The discriminant genes were combined to form signatures that predict the response of tumours to single drugs. Finally, these signatures were combined to produce composite signatures that were shown to predict the response to multidrug regimens in two clinical studies.[26] The main advantage of the '*in vitro* signature' approach is that the learning set contains cell lines, whereas the validation set contains tumours. By definition, this avoids the type of overfitting that occurs when the same data are used to create and test a predictor.

Metagene analysis

We have developed a method that decomposes the gene expression signal into independent groups of co-expressed genes (metagenes) associated with a small number of biological processes that dominate the gene expression profiles of breast tumours in all published microarray studies (epithelial cell type, stroma, T cells, B cells, adipocytes, proliferation, interferon response and hypoxia).[29] The model used to produce the metagenes was not fitted to our own data but to a dataset containing 581 tumours from the Netherlands Cancer Institute and the Erasmus Medical Centre. By limiting the analysis to nine variables, instead of the 61297 probesets on the Affymetrix chip, we drastically reduced the need to correct for multiple testing, which is otherwise a major barrier to identifying significant differences in microarray data. The ability of the metagenes to predict pCR status was assessed in our own data and in published datasets from MDACC and Duke University.[29]

2.2.1.2. *Performance of multigene classifiers of specific chemosensitivity.*

Using the conventional supervised approach, the MDACC group identified a signature that allows correct identification of more than 90% of patients who fail to achieve a pCR (negative predictive value 96%, 95% CI 82–100) but fails to identify almost 50% of patients who do achieve a complete pathological response (positive predictive value 52%, 95% CI 31–73).[23] Gianni and collaborators have reported provocative results through a collaboration between the Istituto Nazionale Tumori (INT) and the MDACC.[22] As a first step, they tested the predictive value of 384 genes whose expression was measured by qRT-PCR in a series of 89 patients treated with neoadjuvant chemotherapy at the INT. They identified 86 genes by conventional supervised analysis predicting for pCR at $p < 0.05$. They then assessed the predictive value of these 86 genes in an independent series of 82 pa-

tients treated at the MDACC. Rather than validate the results using the qRT-PCR technique, they chose instead to perform the validation study on Affymetrix gene expression microarrays. As a result, they were only able to test 79 genes of the genes in the original signature. Of these genes, 24 demonstrated a correlation with pCR at $p < 0.05$, with an expected false discovery rate of four genes. Unfortunately, more detailed performance metrics for this multigene classifier were not given in the publication.

The multigene classifier produced by the Duke team using the *in vitro* signature method was first applied to a data set of 51 breast cancer patients treated with neoadjuvant paclitaxel, 5-fluorouracil, doxorubicin and cyclophosphamide.[26] The accuracy of the multigene classifier was 82% with a 61% positive predictive value and a 94% negative predictive value. In collaboration with the Duke team we recently confirmed the validity of this approach in a subgroup of ER-negative patients included in a recently closed phase III neoadjuvant clinical trial (EORTC 10994/BIG 00-01 study).[27] This trial compares a non-taxane regimen (fluorouracil + epirubicin + cyclophosphamide × 6; FEC arm) with a taxane regimen (docetaxel × 3 followed by epirubicin + docetaxel × 3; T->ET arm). One hundred and twenty-five ER-negative tumours were tested: 66 in the FEC arm (28 pCR) and 59 in the T->ET arm (27 pCR). The Duke team created multigene classifiers specific for the drugs included in each arm of the trial. These regimen-specific multigene classifiers significantly predicted pCR in patients treated in the appropriate arm ($p < 0.0001$). The FEC predictor had a sensitivity of 96% (27/28 patients), specificity 66% (25/38), PPV 68% (27/40) and NPV 96% (25/26). The T->ET predictor had a sensitivity of 93% (25/27), specificity 69% (22/32), PPV 71% (25/35) and NPV 92% (22/24). Analysis of tumour size, grade, nodal status, age and the regimen-specific multigene classifiers showed that the multigene classifiers were the only independent variables predicting pCR at $p < 0.01$.

Using the metagene approach described above, we recently reported that increased stromal gene expression predicts resistance to FEC chemotherapy in a subgroup of patients enrolled in the EORTC 10994 study, with an AUC of 0.68 (95% CI 0.54–0.80; $p = 0.03$).[29] The predictive value of the stromal multigene classifier was successfully validated in two independent cohorts of patients who received chemotherapy,[29] but the accuracy of the prediction was only 65%, so further validation and greater understanding of the underlying resistance mechanism will be required before attempting to exploit the information in the clinic.

2.2.1.3. Interpretation of multigene classifiers of specific chemosensitivity. The number of patients included in the studies in Table 1 ranged from 30 to 171, but the highest numbers in the validation sets were only 51 in the MDACC study,[23] 66 and 59 in the EORTC study[27] and 82 in the INT/MDACC study.[22] These are small numbers so we should expect some of the reported signatures not to withstand rigorous external validation. The MDACC group reported the largest series and the most thorough analysis using conventional supervised techniques.[23] However, it is unlikely that the low positive predictive value of their signatures can be increased by further optimisation of the techniques. As shown by Hess and collaborators, whether one increases the number of tumours in the training set, uses a higher number of probesets in the signature, or uses a different classification algorithm, results are not improved.[23]

In the breast cancer literature there is less data available with the *in vitro* signature approach than with conventional supervised techniques. The positive and negative predictive values obtained with the *in vitro* signature approach look promising[26,27] but need to be confirmed. To address this question, we are now conducting an *in vitro* signature-based study on 200 luminal B cancers from the EORTC 10994/BIG 00-01 trial. The Duke group is continuing to develop the technique and recently published data on 59 patients with advanced epithelial ovarian cancer previously treated with cisplatin.[32] The accuracy of the cisplatin predictor was 83%, with a 78% positive predictive value and a 100% negative predictive value.

It is important to understand what is meant by prediction in these neoadjuvant studies. Only by studying tumours from patients randomised to two treatment regimens is it possible to identify regimen-specific signatures and to distinguish between predictive and prognostic signatures. If patients are not randomised to different treatments (as was the case in all the studies in Table 1 with the exception of the EORTC trial), the only conclusion which can be drawn is that the signature is predictive of better response, and thus perhaps of general chemosensitivity. The result does not allow one to say that the signature predicts sensitivity to a specific class of agents. It is questionable whether this is really clinically useful, since it offers no alternative strategy for patients predicted not to respond to the therapy: it cannot exclude some benefit from the therapy, and does not identify a better one! More data from randomised trials are therefore urgently needed.

The meaning of a multigene classifier is even more complex when determined in the context of an adjuvant study in which all patients are given the same chemotherapy. For example, a recently reported study identified a 14-gene signature which predicts the development of metastatic disease after treatment with FEC, a standard anthracycline-based regimen.[33] However, because of the study design, it is not possible to determine how much of the power of this signature is purely prognostic and how much can be attributed to drug resistance. Data from large adjuvant randomised trials are eagerly awaited in addition to those described above suggesting a role for the 21-gene signature (Oncotype DX™) as a predictor of global chemosensitivity.[11,12]

2.2.2. Genomic and post-transcriptomic analysis

The majority of studies to date have focussed on transcriptomics, probing differences in gene expression at the RNA level. By testing RNA, one interrogates the phenotype of the tumour cell. Proteomics and metabolomics are the logical endpoints of phenotypic analysis, although they have had limited impact so far. The alternative is to look directly at the DNA level for the underlying genetic defects that drive oncogenic growth, in other words, to examine the genotype of the tumour cell. Two genotypes need to be considered, that of the germ line and that of the tumour cell. Germ line studies are used to address questions about drug metabolism, for example, to identify single nucleotide polymorphisms (SNPs) that affect

prodrug activation. More commonly, genotypic studies have focused on changes in the tumour cells themselves. Several groups have reported the results of array-based comparative genomic hybridisation (CGH) studies in breast cancer.[34–38] We have only identified two studies that tried to predict chemotherapy response using genomic profiles.[37,38] Analysing the DNA changes (either losses or gains) in tumours of 44 and 106 patients respectively, all of whom had been treated with neoadjuvant chemotherapy, the authors were unable to predict clinical or pathological response. We will suggest an explanation for this negative result in Section 4.2.

3. Prospective studies

As far as predictive multigene classifiers for global chemosensitivity are concerned, we will summarise two large ongoing phase III studies (MINDACT and TAILORx) that use multigene classifiers both to identify patients who do not need chemotherapy and to identify patients who are most likely to benefit from chemotherapy. Finally, we will discuss the design of a phase III trial under development whose aim is solely to validate predictive multigene classifiers for specific chemotherapies.

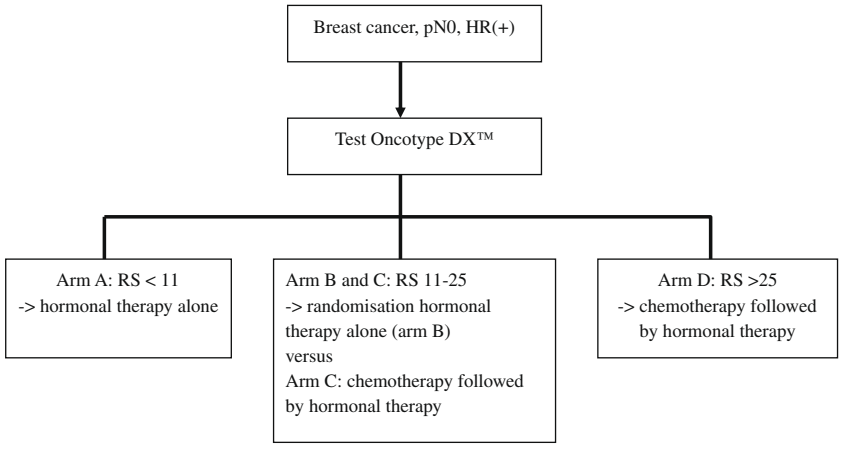
3.1. *Predictive multigene classifiers for general chemosensitivity*

3.1.1. *TAILORx trial*

The Breast Cancer Intergroup (TBCI) is conducting a prospective trial called TAILORx (Trial Assigning Individualized Options for Treatment; Fig. 1) in the USA in the context of the Program for the Assessment of Cancer Clinical Tests (PACCT).[39] Patients with node-negative hormone receptor-positive tumours are eligible for this trial. Those with a low recurrence score (RS < 11) are treated with endocrine therapy alone. Those with a high recurrence score (RS > 25) are treated with endocrine therapy plus chemotherapy. Those with an intermediate recurrence score (RS = 11–25) are randomised to receive either endocrine therapy alone or endocrine therapy plus chemotherapy. In this trial the levels of recurrence score defining the three groups (low RS group, intermediate RS group, and high RS group) are slightly different from the scores defined in the retrospective trials previously published. The cut-offs were changed to decrease the risk of under treatment in the intermediate and high recurrence score groups. The new cut-offs were validated by analysing the results of the NSABP B-20 study: the benefit of chemotherapy at the new cut-offs was similar in the intermediate and high recurrence score groups and the risk of relapse was less than or equal to 5% in the low and intermediate recurrence score groups.[39] The statistical design in the intermediate recurrence score group uses a non-inferiority rule and has sufficient power to detect a 3% difference or more between the two treatment arms. The accrual in this trial is going well and should be completed in 2011.

3.1.2. *MINDACT trial*

The EORTC and the Breast International Group (BIG) are conducting a prospective trial called MINDACT (Microarray in Node-Negative Disease May Avoid Chemotherapy) to validate the prognostic value of the 70-gene signature (Mammaprint™, Fig. 2). The trial will also evaluate the predictive value of the signature in the so-called 'discordant group'.[40] Patients with invasive breast cancer, up to three positive nodes and any hormone receptor status are eligible for this trial. Risk scores are calculated with both the Mammaprint signature and the 'Adjuvant!' algorithm,[41] which is based on conventional clinical risk factors. Patients with high risk by both scores receive chemotherapy, whereas patients with low risk by both scores receive hormonal therapy; in these groups the intention is simply to validate the prognostic value of the scores. When the scores disagree, patients are randomised to have risk (and thus chemotherapy) assigned by either the Adjuvant! score or the Mammaprint score, as shown in Fig. 2. In the 'discordant group' the intention is to reduce the number of patients receiving unnecessary treatment.



Abbreviations
HR (+): hormone-receptor-positive;
RS: recurrence score

Fig. 1 – TAILORx design.

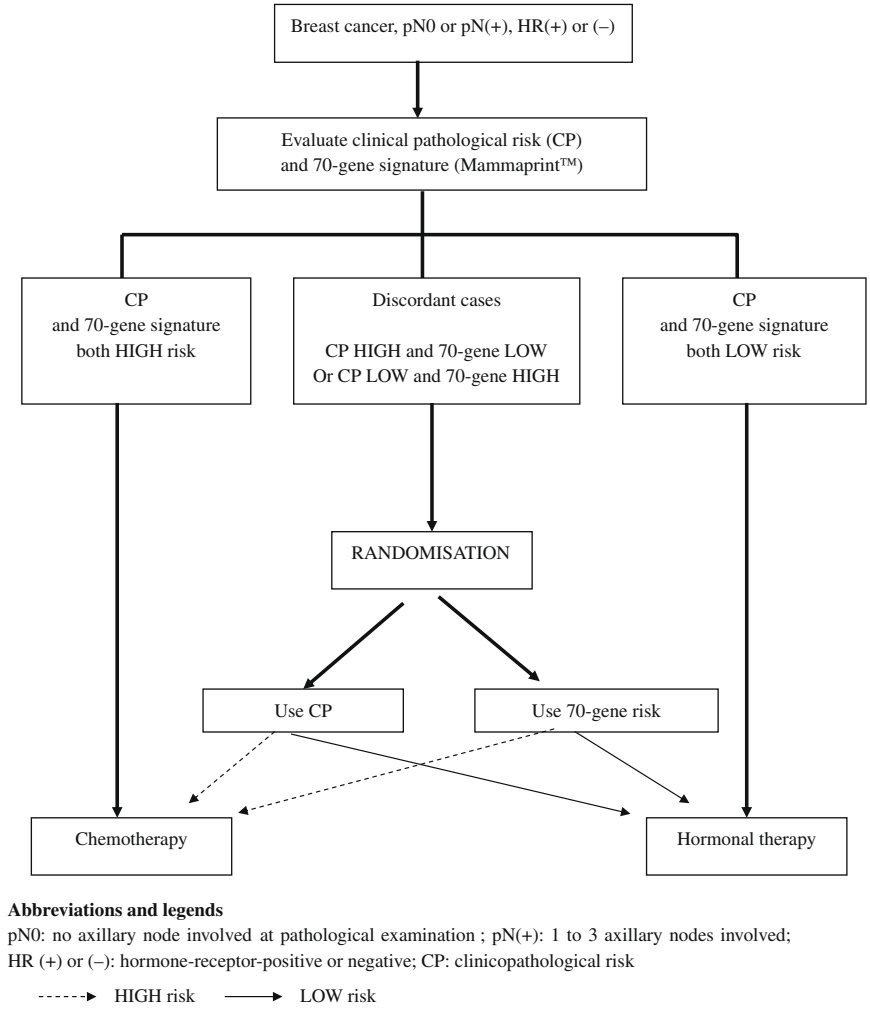


Fig. 2 – MINDACT design.

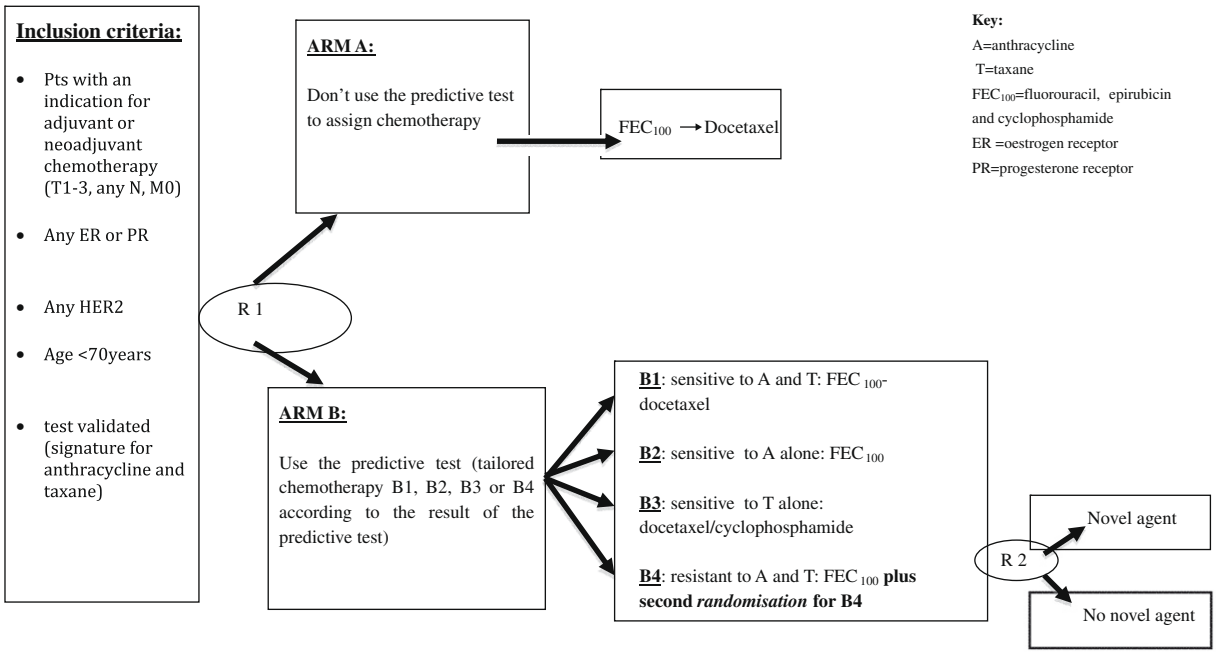


Fig. 3 – Example of a possible study design testing the chemotherapy-specific signature-based approach.

3.2. *Predictive multigene classifiers for specific chemosensitivity*

The ultimate research question is whether selection of chemotherapy using multigene classifiers leads to fewer relapses. Several randomised phase II trials are ongoing in the neoadjuvant setting aiming to improve the pathological complete response (pCR) rate using multigene classifiers. However, improving pCR rates does not always lead to better survival, and therefore only a phase III study using a marker-based design[42] will generate level-of-evidence 1[43] confirming that this strategy improves patient outcomes. In this approach, a patient is assigned either to receive treatment according to the multigene classifier, or to receive treatment independently of the multigene classifier. The predictive value of the multigene classifier can be determined by comparing the outcomes in the two groups.

An example of a trial design testing this chemotherapy-specific signature-based approach is shown in Fig. 3, as recently discussed within the BIG. We aim to conduct this trial in patients with invasive breast cancer suitable for adjuvant or neoadjuvant chemotherapy, with any hormone receptor status and any *HER2* status. Considering the promising results obtained with the '*in vitro* signature' algorithm[27] we intend to use this algorithm to predict response. For the sake of simplicity we plan to use only the multigene classifiers for the single agents doxorubicin and docetaxel, rather than the full polychemotherapy response predictors. This phase III trial will have two objectives. Firstly, to improve the outcome for all patients by selecting the most appropriate treatment for patients whose tumours are predicted to be sensitive to anthracycline or docetaxel or both. As shown in Fig. 4, we will have four groups (B1: sensitive to both anthracycline and docetaxel; B2: sensitive to anthracycline only; B3: sensitive to docetaxel only; B4: resistant to both anthracycline and docetaxel). Secondly, to improve the outcome for patients whose tumours are predicted to be resistant to both anthracyclines and taxanes (group B4 in the figure) by adding a novel agent to existing chemotherapy in this subgroup. Several approaches are currently under discussion for selection of the treatment in group B4. One option would be to select the treatment based on the molecular classification of breast tumour type: for example, randomisation between a new agent 'A' or no new agent for patients in the Basal-like tumour class and randomisation between a new agent 'B' or no new agent in the Luminal B tumour class. Another option would be to select the new treatment based on oncogenic pathway deregulation signatures. Several gene expression signatures associated with specific oncogenic pathway deregulation (*p53, PI3K, Ras, MYC, Src, E2F3* and *b-catenin*) have been reported in the literature,[44–46] and studies are ongoing to confirm the robustness of the signatures. Pathway signatures would allow selection of patients who are most likely to benefit from new targeted therapies. A third option would be to use genomic data ('genotype' in Fig. 4), which is easier to quantify and does not require such complex algorithms for interpretation.[36] This approach is immediately appealing to clinicians since drugs targeting these pathways are under development (for example, *Src* inhibitors and *PIK3CA* inhibitors). This approach should also appeal to the pharmaceutical industry since the selection of patients with a genomic classifier capable of predicting the response to the new therapy would reduce the number of patients needed to demonstrate a therapeutic effect.[47]

There are at least three prerequisites before one can use a multigene classifier to select a specific chemotherapy regimen for an individual patient in the clinic: (1) a laboratory that can generate reliable and reproducible multigene classifiers sufficiently fast enough so that the information is available before treatment begins; (2) a tissue preservation technique that lends itself to routine clinical use (fresh frozen, RNAlater or paraffin embedded samples); and (3) a validated treatment strategy based on the multigene classifier. The Pilot MINDACT[48] and MINDACT trials have, to some extent, already addressed the first issue but only in the sense

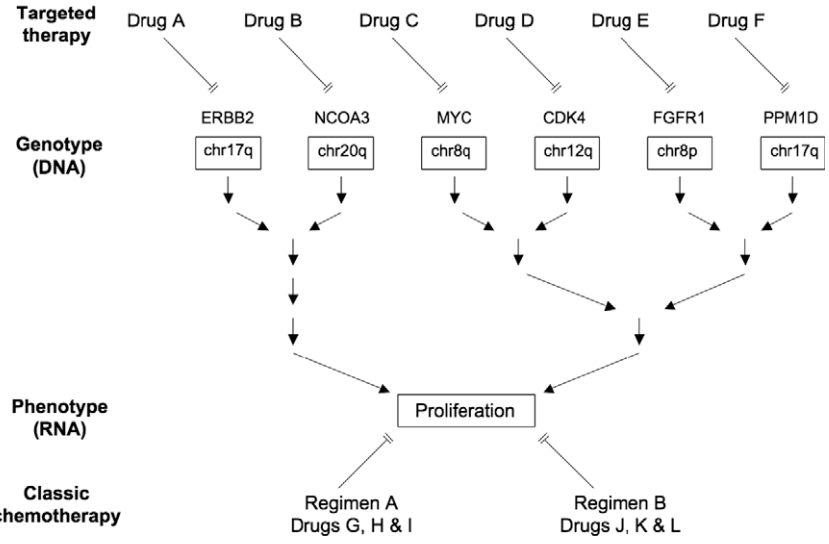


Fig. 4 – Cause versus effect. Genotype predicts the response to targeted therapy whereas phenotype predicts the response to classic chemotherapy. Note that the arrows are for illustration purposes only and do not reflect the full complexity of the pathways involved.

of a single high/low molecular risk categorisation. In order to address the other issues, plans are underway to run a 300 patient feasibility study in France and the UK (called 'Pilot PREDICT'). It is anticipated that this study will open to accrual before the end of 2009, and recruit patients over a 6 month period.

4. Did we get more insight into the biological understanding of resistance to chemotherapy and the way to overcome this resistance? How to identify new targets?

4.1. *Identification of new drug resistance mechanisms?*

In the late 1990s, when the current wave of clinical microarray studies was initiated, there was a strong expectation that microarray analysis of tumours would lead to the identification of new drug resistance mechanisms. For example, it was plausible to imagine that increased expression of a previously untested microtubule binding protein might account for a substantial part of the variation in the response to taxanes. Sadly, it looks increasingly likely that few genuinely new drug resistance mechanisms will be brought to light by microarray studies on gene expression in tumours.

4.2. *How to identify new targets?*

Why has the genomic revolution in the form of expression profiles generally failed to deliver new therapeutic targets or easy response prediction? There are probably many reasons, including the lack of correlation between mRNA and protein levels for some genes, and the fact that drug metabolism depends on activities outside the tumour, such as hepatic cytochrome P450 activity. However, it is also fundamentally important to distinguish cause from effect. Fig. 4 explains why we should not be surprised by the failure to identify new targets through expression profiles. Many different lesions at the DNA level (causes) converge to create the malignant phenotype (effect). A key phenotype of malignant cells is their high growth rate, the therapeutic potential of which is exploited by classical chemotherapy. In contrast, single oncogenic mutations rarely have major effects on chemotherapy response because they are far upstream of the targets of classic chemotherapy, such as the DNA replication process or the microtubules of the mitotic spindle. This does not mean genomic changes lack predictive information, they simply contain a different sort of information. Specifically, they contain information about the response to therapies targeted at those individual changes. The classic example is *ERBB2* amplification, which predicts response to trastuzumab (drug A in Fig. 4). The challenge for the pharmaceutical industry and the oncology community is to develop drugs B–F. The tools are now in place to perform large-scale genotyping of tumours at reasonable cost to identify new targets, for example, by hybridisation of tumour DNA to microarrays (CGH arrays, bead arrays and SNP chips) or whole genome sequencing (Illumina SOLEXA, ABI Solid or Roche 454 platforms), and to develop the drugs to inhibit them, for example, by high throughput screening of small molecules and structure-based drug design.

5. Conclusion

Are predictive signatures ready for everyday clinical use? With regard to classifiers of global chemosensitivity, two large phase III trials are ongoing (TAILORx and MINDACT), and we can expect an answer in the coming years. Regarding classifiers of specific chemosensitivity, the priority is to demonstrate in a phase III trial that use of these classifiers leads to improved disease-free survival of the whole population of early breast cancer patients. We have presented in this review a possible design for such a trial. The trial will be expensive with current technologies, but it represents an important step towards the development of personalised medicine. If the approach works, it will reduce the current split between prognostic and predictive approaches: the clinician will be able to discuss with the patient both a general measure of the risk of relapse, and a specific measure that predicts which (if any) chemotherapeutic regimen can reduce the risk sufficiently to warrant the toxicity of that particular treatment. In addition, other priorities for future research should be to validate oncogenic pathway signatures as tools to predict drug response, to incorporate information on drug metabolism into predictive signatures (e.g. by genome wide SNP analysis) and to identify new drug targets through large-scale genotyping of tumours.

Conflict of interest statement

None declared.

Acknowledgements

We would like to thank John Bartlett, Judith Bliss and Nick Murray for helpful discussions about the proposed PREDICT projects.

REFERENCES

1. Dowsett M, Goldhirsch A, Hayes DF, Senn HJ, Wood W, Viale G. International web-based consultation on priorities for translational breast cancer research. *Breast Cancer Res* 2007;**9**(6):R81.
2. Henderson IC, Berry DA, Demetri GD, et al. Improved outcomes from adding sequential paclitaxel but not from escalating doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer. *J Clin Oncol* 2003;**21**(6):976–83.
3. Hayes DF, Thor AD, Dressler LG, et al. HER2 and response to paclitaxel in node-positive breast cancer. *N Engl J Med* 2007;**357**(15):1496–506.
4. Berry DA, Cirrincione C, Henderson IC, et al. Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA* 2006;**295**(14):1658–67.
5. De Laurentiis M, Cancello G, D'Agostino D, et al. Taxane-based combinations as adjuvant chemotherapy of early breast cancer: a meta-analysis of randomized trials. *J Clin Oncol* 2008;**26**(1):44–53.
6. Andre F, Broglio K, Roche H, et al. Estrogen receptor expression and efficacy of docetaxel-containing adjuvant

chemotherapy in patients with node-positive breast cancer: results from a pooled analysis. *J Clin Oncol* 2008;**26**(16):2636–43.
7. Desmedt C, Ruiz-Garcia E, Andre F. Gene expression predictors in breast cancer: current status, limitations and perspectives. *Eur J Cancer* 2008;**44**(18):2714–20.
8. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 2009;**360**(8):790–800.
9. Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol* 2005;**23**(29):7332–41.
10. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;**351**(27):2817–26.
11. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;**24**(23):3726–34.
12. Albain K, Barlow W, Shak S, et al. prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal, node positive, ER-positive breast cancer (S8814,INT01100). *Breast Cancer Res Treat* 2007:abstract 10.
13. van't Veer L, Dai H, van de Vijver M. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;**415**(4):530–5.
14. van de Vijver MJ, He YD, van't Veer L, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;**347**(25):1999–2009.
15. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006;**98**(17):1183–92.
16. Mook S, Schmidt MK, Viale G, 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*; July 27 [Epub ahead of print]. doi:10.1007/s10549-008-0130-2.
17. Straver ME, Glas AM, Hannemann J, et al. The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res Treat*; Feb 13 [Epub ahead of print]. doi:10.1007/s10549-009-0333-1.
18. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003;**362**(9381):362–9.
19. Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004;**22**(12):2284–93.
20. Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. *J Clin Oncol* 2005;**23**(3):422–31.
21. Hannemann J, Oosterkamp HM, Bosch CA, et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005;**23**(15):3331–42.
22. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol* 2005;**23**(29):7265–77.
23. Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006;**24**(26):4236–44.
24. Thuerigen O, Schneeweiss A, Toedt G, et al. Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer. *J Clin Oncol* 2006;**24**(12):1839–45.
25. Cleator S, Tsimelzon A, Ashworth A, et al. Gene expression patterns for doxorubicin (Adriamycin) and cyclophosphamide (cytoxan) (AC) response and resistance. *Breast Cancer Res Treat* 2006;**95**(3):229–33.
26. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006;**12**(11):1294–300.
27. Bonnefoi H, Potti A, Delorenzi M, et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. *Lancet Oncol* 2007;**8**(12):1071–8.
28. Chang JC, Makris A, Gutierrez MC, et al. Gene expression patterns in formalin-fixed, paraffin-embedded core biopsies predict docetaxel chemosensitivity in breast cancer patients. *Breast Cancer Res Treat* 2008;**108**(2):233–40.
29. Farmer P, Bonnefoi H, Anderle P, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med* 2009;**15**(1):68–74.
30. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;**365**(9460):671–9.
31. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;**95**(1):14–8.
32. Hsu DS, Balakumaran BS, Acharya CR, et al. Pharmacogenomic strategies provide a rational approach to the treatment of cisplatin-resistant patients with advanced cancer. *J Clin Oncol* 2007;**25**(28):4350–7.
33. Campone M, Campion L, Roche H, et al. Prediction of metastatic relapse in node-positive breast cancer: establishment of a clinicogenomic model after FEC100 adjuvant regimen. *Breast Cancer Res Treat* 2008;**109**(3):491–501.
34. Naylor TL, Greshock J, Wang Y, et al. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. *Breast Cancer Res* 2005;**7**(6):R1186–98.
35. Nessling M, Richter K, Schwaenen C, et al. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. *Cancer Res* 2005;**65**(2):439–47.
36. Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 2006;**10**(6):529–41.
37. Pierga JY, Reis-Filho JS, Cleator SJ, et al. Microarray-based comparative genomic hybridisation of breast cancer patients receiving neoadjuvant chemotherapy. *Br J Cancer* 2007;**96**(2):341–51.
38. Andre F, Job B, Dessen P, et al. Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. *Clin Cancer Res* 2009;**15**(2):441–51.
39. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J Clin Oncol* 2008;**26**(5):721–8.
40. 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**(5):729–35.
41. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001;**19**(4):980–91.
42. Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol* 2005;**23**(9):2020–7.
43. Hayes DF, Bast RC, Desch CE, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* 1996;**88**(20):1456–66.

44. Miller LD, Smeds J, George J, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 2005;**102**(38):13550–5.
45. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;**439**(7074):353–7.
46. Salter KH, Acharya CR, Walters KS, et al. An integrated approach to the prediction of chemotherapeutic response in patients with breast cancer. *PLoS ONE* 2008;**3**(4):e1908.
47. Simon R, Maitournam A. Evaluating the efficiency of targeted designs for randomized clinical trials. *Clin Cancer Res* 2004;**10**(20):6759–63.
48. Mook S, Bonnefoi H, Pruneri G, et al. Daily clinical practice of fresh tumour tissue freezing and gene expression profiling; logistics pilot study preceding the MINDACT trial. *Eur J Cancer* 2009;**45**(7):1201–8.