# REVIEW

# The impact of expression profiling on prognostic and predictive testing in breast cancer

## J S Reis-Filho, C Westbury, J-Y Pierga

.........................................................................................................................

Expression profiling has been extensively applied to the study of breast cancer and undoubtedly is changing the way breast cancer is perceived. Over the past few years, several groups have described prognostic "signatures" (gene lists) that are purported to be more accurate prognostic factors than well established clinical and pathological features. In addition, cDNA and oligonucleotide microarrays have also been used to devise predictive "signatures" in the setting of neoadjuvant chemotherapy setting. However, it seems that the enthusiasm with this new technology has led most of us to turn a blind eye to some serious methodological problems which are evident in landmark papers on breast cancer expression profiling. These issues include small and biased cohorts of patients, inappropriate statistical analysis and lack of thorough validation of the technology. In this review, we critically revisit the most relevant cDNA microarray studies on breast cancer prognosis and prediction published to date. Although the results are promising, further optimisation and standardisation of the technique and properly designed clinical trials are required before microarrays can reliably be used as tools for clinical decision making.

.............................................................................

B reast cancer is a heterogeneous disease which encompasses several entities with distinct prognosis. Although a comprehensive breast cancer morphological taxonomy has been developed and usefully applied to patient management, it has become clear that tumours classified under the same umbrella descriptive term may have distinct underlying biological features and clinical behaviour.[1–7]

Since the advent of antibodies that can be applied to formalin fixed, paraffin embedded tumour sections, we have seen the rise and fall of several prognostic and predictive biological markers.[5 8–10] Until the late 1990s and the boom of high throughput methodologies, the main approach used for identifying prognostically significant groups consisted of testing one or a few markers in a cohort of patients, usually retrospectively. Although a plethora of studies using this approach have been published, only hormone receptors (progesterone and oestrogen receptors (ER))[4] and HER-2[11] have been translated into clinical markers for routine prognostic use in breast cancer management.

With the development of tailored therapies targeting specific molecular markers, ER and HER2/neu have also become important predictive factors, as patients with ER positive tumours may benefit from being treated with selective oestrogen receptor modulators (SERM) and aromatase inhibitors,[4] whereas patients with HER2/neu positive tumours have been shown to experience a significant survival advantage when treated with humanised monoclonal antibodies against HER2/neu.[11] Hence, future identification of novel prognostic and predictive factors has become imperative to further individualise therapy of breast cancer.

The survival benefit of adjuvant systemic therapy has been well documented for over 30 years, and it is becoming increasingly evident that distinct therapies have differential benefit in specific subgroups.[12 13] A paradigmatic example is the aforementioned benefit of humanised antibodies against HER2 for the treatment of patients with HER2 positive tumours.[11]

Since the completion of the human genome sequencing and the development of high throughput techniques, analysis of the expression of thousands of genes in a given tumour has become possible.[14 15] This approach has not only furthered our understanding of breast cancer taxonomy,[16–19] but also provided a number of "signatures" (collection of genes that taken together can classify tumours into distinct groups, sometimes with prognostic or predictive implications) that have been reported to be more effective than standard prognostic and predictive factors.[5 16–32]

Not surprisingly, apart from a few exceptions, most pathologists have been reluctant to deal with this new technology and feel that their role in "guiding the surgeon's hand" would be in jeopardy.[7] This has been aggravated by the attitude of some scientists and clinicians, who have deemed current pathology methods as unsophisticated and obsolete, and compared them to some ritualistic practices of primitive tribes.[33] Although some may perceive the pattern recognition methods of diagnostic histopathology as an obsolete analytical method, and haematoxylin and eosin stained slides as very rudimentary tools, the histological appearance of a given tumour may be considered the final product of the orchestrated interaction between different classes of genes and proteins (growth

.................................................................

See end of article for authors' affiliations

......................

Correspondence to:
Dr J S Reis-Filho, The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK; jorgerf@icr.ac.uk

**Abbreviations:** AC, doxorubicin cyclophosphamide; AD, doxorubicin docetaxel; cDNA, complementary DNA; CSR, core serum response; ER, oestrogen receptor; Q-RT-PCR, quantitative reverse transcriptase polymerase chain reaction; SERM, selective oestrogen receptor modulator

factors and their receptors, cell cycle regulators, transcription factors, apoptosis inhibitors and promoters, matrix proteins) and different cell types.

We do not argue that, in the current era of tailored therapy, expression profiling is ready to guide our decision making in management of breast cancer patients.[5 22 23 25 26 28–30 32 34] Development of technology in this area has been rapid, but there are several issues yet to be resolved.[7 35] There are several excellent reviews on the impact of expression profiling on the management of breast cancer patients.[2 5 24 36 37] Although the contribution of this technology to our understanding of breast cancer biology and treatment is undeniable, the initial enthusiasm has given way to a more critical analysis of the conclusions we can draw from expression profiling studies.[35 36 38] In this review, we aim to offer a fair and balanced view of the contribution of high throughput technology to the treatment of breast cancer patients.

## GENE EXPRESSION PROFILING: WHAT IS IT?

Gene expression profiling refers to any method that can analyse the expression of selected genes in selected samples. For this particular review, we have focused only on complementary DNA (cDNA) and oligonucleotide arrays or chips published before 1 June 2005.

cDNA microarrays are composed of a collection of DNA segments spotted, in a grid arrangement, onto a solid support (''chip''; glass slide or fibrous mesh membrane). The microarray spots serve as hybridisation targets for cDNA, representing messenger RNA extracted from tissue samples or cell lysates. cDNA is synthesised by reverse transcription of the extracted RNA from the test sample and a reference sample is also prepared in this way.[14] The test sample and reference sample are differentially labelled with fluorophores, and then combined and hybridised to the array under controlled conditions. After stringent washes to remove non-specific hybridisation, reference and sample cDNA will only hybridise to complementary sequences on the arrays. The ratio of gene expression between test and reference samples for a given gene determines the colour and intensity of each microarray spot, which can be measured. Starting material from patient samples (core needle biopsies) is often limited, and amplification methods using in vitro transcription have been applied to yield sufficient quantities of material to array. In this case, amplified RNA from test and reference samples is labelled and hybridised to the chip.

Oligonucleotide microarrays or ''chips'' are composed of oligonucleotides synthesised in situ on a solid substrate. They follow the same principles of cDNA microarrays, although some ''chips'' do not require the concurrent hybridisation of a reference sample, and the expression levels are defined according to mathematical algorithms rather than a direct ratio between tumour and reference mRNA.

Vast amounts of data representing expression levels of many thousand genes for any given sample can be generated by these techniques, and increasingly sophisticated analytical methods are being developed to process and make sense of these data.

## ARRAYS AND REPRODUCIBILITY: ARE WE READY TO USE THEM AS CLINICAL PATHOLOGY TESTS?

Although array technology is an evolving field, and huge sums of money have been involved in creating new and more comprehensive platforms for gene expression profiling, quality control and reproducible analysis systems have yet to be fully defined. Several publications have addressed issues of quality control relating to the microarray platform and aspects of data analysis; however, very little has been reported on quality control of tissue handling.

Recently, a series of studies evaluated the reproducibility of expression profiling experiments.[39–48] Initial results showed an exceedingly poor correlation between the results obtained with different platforms, suggesting that data obtained with cDNA microarrays, and oligonucleotide chips would not be comparable.[39 45–47] Further and more controlled experiments were carried out, comparing intralaboratory and interlaboratory reproducibility, and although a perfect agreement is yet to be achieved, the results have been more encouraging.[41–44 48] Some aspects of expression profiling have been analysed in depth, such as probe size, sequence homology, and hybridisation protocols; however, the changes in the expression profile associated with the tissue handling process, in particular the initial retrieval of the clinical specimen and subsequent transport and storage before processing into frozen blocks, have not been addressed in detail. RNA is inherently unstable and, in addition, rapid changes in gene expression may occur as a result of insults caused by tissue handling and iatrogenic ischaemia.[49] Therefore, there is potential for marked variability in results of expression profiling purely due to technical issues relating to specimen processing.

A step towards standardisation of initial tissue collection has been achieved with the use of fixatives that can preserve RNA without significantly causing loss of tissue morphological detail. Development of stabilisation media such as RNALater[50] has allowed rapid tissue processing and some potential for standardisation of the tissue collection process. The relative simplicity of immersing a clinical specimen into RNALater in comparison with traditional methods such as snap freezing in liquid nitrogen makes this stabilisation process easier to standardise. In addition, variation in expression profile in tissue stored at room temperature in RNALater between 24 and 72 hours does not appear to alter the tissue expression profile significantly,[50] which again allows some leeway in the transport of samples to the laboratory for further storage and processing.

Not only differences in initial handling of the specimen but also subsequent processing of extracted RNA may contribute to changes in RNA integrity. Furthermore, different authors have shown that variability in RNA integrity can significantly bias gene expression data.[51 52] For example, Imbeaud et al[52] showed, using quantitative reverse transcriptase PCR (Q-RT-PCR) that expression levels of housekeeping genes varied as much as sevenfold when analysing RNA of different integrity from the same source. These factors must certainly be taken into account when considering the robustness of expression profiling as a diagnostic test in routine clinical practice.

## EXPRESSION PROFILING AND EXPERIMENTAL DESIGN

Gene expression profiling studies may focus on unravelling mechanisms in vitro, such as determining drug action or the effect of a given knocked in or knocked out gene in cell lines. Studies using clinical samples allow the analysis of in vivo expression profiles and may also determine mechanistic factors such as disease pathogenesis or response to therapy.[2 5 16–19 21–23 25 26 28–30 32 53 54] Most of the studies applying gene expression profiling to breast cancer can be classified into three groups: (a) class comparison, (b) class prediction, and (c) class discovery (see Simon et al for a review[55]).

Class comparison is the analysis of gene expression in groups of specimens, which are defined by other methods (histopathological features). This type of analysis aims to identify genes that are differentially expressed between the different categories or classes. A good example is the study by Korkola et al,[56] which was a comprehensive comparison between invasive lobular and ductal carcinomas of the breast. In that study, the authors not only confirmed that expression of E-cadherin is relatively lower in lobular carcinomas compared with ductal carcinomas, but also identified other genes that are differentially expressed in these two types of

breast carcinomas (survivin, cathepsin B, TPI1, SPRY1, SCYA14, TFAP2B, and thrombospondin 4).[56]

Class prediction studies also involve the comparison of the expression profile of predefined groups; however, the major aim is to develop a gene expression based function (also known as predictor or signature) that can accurately predict the class membership of new samples solely on the basis of the predictor. Class prediction studies attempting to build predictors of prognosis and response to chemotherapy in breast cancer abound.[2 5 20 22 26 28 29 31 32 53 57] Although the signatures produced by different research groups appear to differ in terms of the genes included, recent studies have shown that these can be used in a complementary fashion (see Chang et al[23] and below).

Class discovery concerns the identification of new classes or groups, regardless of other features. The primary endpoint of these studies is not to compare expression profiles with known features, but to develop a new taxonomy for a given disease. The seminal study by Perou et al,[18] is a good example of this approach, where breast carcinomas were classified into four main groups (described below), based upon the similarity in gene expression to normal cell counterparts.

Each of these types of study, applying gene expression profiling to clinical cases, have several limitations. In a recent review, Simon et al[55] emphasised the importance of design, statistical analysis and validation in microarray studies. As pointed out by these authors, if inappropriate statistical methods or validation sets are used, the conclusions may be applicable only to the dataset used in a given study. Simon et al[55] have also pointed out that unsupervised hierarchical clustering analysis (a mathematical method that has been extensively used in the last few years to identify group samples with similar expression profiles together while separating samples with distinct expression profiles) is not suitable for all studies. In fact, this method is best applied only to class discovery studies. For class comparison and class prediction analyses, supervised analyses seem to be the most appropriate option.[55]

## MICROARRAYS AS PROGNOSTIC FACTORS: A SIGNATURE TO RULE THEM ALL

Currently used prognostic and predictive factors, derived from clinical parameters, histopathology, and immunohisto-chemical markers, have been successfully used for managing patients with breast cancer.[3 4 8] Retrospective and prospective studies have shown that a high proportion of node negative patients undergo systemic chemotherapy because the current methods cannot accurately determine the risk of recurrence or relapse for a given patient.[5 28 29 32] It has become clear that only a minority of these patients will develop a recurrence, therefore more refined predictors of relapse and survival are necessary to avoid overtreating these patients with unnecessary toxic systemic treatment.[2 5 23 28–30 32]

Patients with similar clinical and pathological features have been reported to show distinct outcomes, suggesting the existence of additional underlying molecular features that determine the tumour's behaviour and the possibility that new undiscovered molecular subclasses exist. Progress in unravelling these molecular differences has been made by the use of gene expression profiling studies in class discovery and prediction.

In the seminal studies published by the Perou et al[18] and Sorlie et al,[16 17] it was demonstrated that breast carcinomas can be classified according to the similarity between the genetic profiles of cancer cells and their normal counterparts. Using this approach, the authors developed an ''intrinsic gene set'', and using hierarchical clustering analysis, tumours were classified into four main groups: (a) luminal cell-like (tumours that express oestrogen receptor and show profiles

similar to those of normal luminal cells); (b) basal cell-like (hormone receptor negative tumours that express genes usually expressed by basal/myoepithelial cells); (c) Erb-B2 (HER-2) (tumours that consistently overexpress HER-2 and are known to harbour HER-2 amplification); and (d) normal breast-like group (which consistently clusters together with normal breast samples and fibroadenomas).[16 18] Interestingly, the authors expanded the series of tumours analysed and showed that the luminal group could be subdivided into three groups: luminal A, B, and C.[16–18]

When comparing the prognosis of tumours of the different groups, it was shown that basal-like or HER-2 tumours showed a more aggressive clinical behaviour, whereas luminal A tumours were associated with an excellent prognosis.[16] Although the approach is very appealing and makes biological sense, prognostic information was available for only a handful of patients. It is puzzling how clinicians and scientists have taken these data so enthusiastically, as there has been no proper validation of the impact of such classification in a large cohort of patients or prospective clinical trials. In addition, this classification approach cannot be used prospectively to classify new samples, as the dendrograms of hierarchical clustering analysis are re-organised when a new sample is added.[5]

Demonstration that luminal-like tumours further subdivide into three categories with different prognoses (luminal A, B, and C) shows how powerful this type of approach can be in class discovery. Rather than superseding classical pathology, these data provide an important potential link back to the pathologist. It has been shown that at least some of these tumour subtypes could be distinguished using a combination of morphology and new immunohistochemical markers identified from the published expression profiles.[58 59] In fact, recent studies using immunohistochemical markers to define the four main groups have provided evidence to support this new breast cancer taxonomy.[58 59] This type of approach for taking new diagnostic entities discovered using genomic techniques into routine diagnostic and clinical practice is still more realistic for the foreseeable future than is widespread expression profiling.[7] A further advantage of using conventional pathological techniques is that with archival samples, a much larger cohort of patients with better defined clinical outcome data can be combined and analysed using tissue microarrays for cross validation.

More recently, Van't Veer et al,[29] through the analysis of 78 tumours using an oligonucleotide array containing 24 479 genes, have developed a 70 gene signature that could classify young (<55 years old) lymph node negative patients into two groups: good prognosis (no recurrence in 5 years of follow up) and poor prognosis (recurrence/metastasis within 5 years of follow up). The classifier optimised for maximum accuracy correctly predicted the outcome in 65 of 78 tumours (83%). When optimised for maximum sensitivity (that is, for the lowest error rate in classifying patients with poor prognosis), the signature correctly classified 31 out of 34 patients in the poor prognostic group. The authors also compared their predictor with the NIH[3] and St. Gallen[4] consensus criteria. Although the 70 gene signature outperformed the latter two in sparing patients from unnecessary chemotherapy, it showed a slightly lower sensitivity for classifying poor prognosis patients (91% v 97% and 94%, respectively).[29]

Using the same 70 gene signature, Van de Vijver et al[28] extended this analysis to a cohort of 234 cases, but this time including patients with stage I and II breast cancer and both node positive and node negative disease. Although in this study, the authors claim that the 70 gene signature outperformed St Gallen[60] and NIH[3] consensus criteria for both low risk and high risk patients, all patients were treated

according to clinical and pathological features and not on the 70 gene signature. Therefore, it is not yet clear that the signature will be more accurate than other methods when patient management is decided solely on the expression profiling data. In addition, if this 70 gene signature is to be incorporated into clinical practice, clinicians may face the situation where the patient has a clinical pathological criteria for poor prognosis and a good gene signature. In this hypothetical situation, if chemotherapy was not offered based upon the good signature, and the tumour recurred, the oncologist could face litigation. Having said that, these issues will be addressed in the Microarray for Node Negative Disease May Avoid Chemotherapy (MINDACT) trial, which is scheduled to begin in September 2005.[61] This prospective randomised trial will test the efficacy of the 70 gene signature and compare it to clinical criteria based on the Adjuvant! Online program (www.adjuvantonline.com/).[62] The power calculation has shown that at least 6000 patients will need to be entered into this trial. Although the many contentious issues may be clarified, the trial could potentially take a long time.

Huang et al,[31] using a complex mathematical method, developed a metagene prognostic signature that could classify individual breast tumours by their likelihood of having associated lymph node metastases at diagnosis and 3 year recurrence risk. Metagenes are not actual genes, but features that encompass much of the discriminatory information in a given cluster of genes. Using mathematical algorithms, lymph node negative tumours could be distinguished from lymph node positive tumours with these metagenes. After training, the system could then classify unknown samples. This type of theoretical model, combining the information from a multitude of genes to provide an accurate molecular classification for difficult clinical problems, is promising. However, this study has several shortcomings, namely the use of lymph node involvement as a surrogate marker for poor prognosis and the lack of a formal validation set.[31][63] In addition, this mind boggling mathematical approach may hinder the characterisation of biological features associated with each group.

Based upon the similarities between wound healing and cancer, Chang et al[21] studied the expression profile of fibroblasts in response to serum exposure, using cDNA microarrays containing approximately 36 000 different genes. The transcriptomic features of fibroblasts grown in the presence of serum appears to reflect the multifaceted role of fibroblasts in wound healing. Analysis of the transcriptomic patterns demonstrated that fibroblasts from different sites have distinctly different gene expression programmed; however, 677 genes were concordantly induced in response to serum in fibroblasts from different sites. Knowing that proliferation is one of the biological phenomena induced by serum, the authors attempted to exclude genes directly related to cell proliferation, resulting in a fibroblast core serum response (CSR) signature comprising 512 serum responsive and cell cycle independent genes. Interestingly, the authors observed that a proportion of breast, lung, and gastric carcinomas express the wound response signature, and that these tumours proved to have a poor overall survival and a high proclivity for metastatic spread.[21] Chang et al[21] then applied their signature to the same patients used in the study of Van de Vijver et al,[28][29] and observed that tumours with the wound response signature showed a decreased probability of being free from distant metastasis and a shorter overall survival when compared with tumours with a quiescent signature. In addition, the CSR signature outperformed the St Gallen[4] and NIH[3] consensus criteria in a cohort of 185 patients who had never received chemotherapy.

In an attempt to combine different gene signatures for clinical decision making, Chang et al[23] developed an approach to integrate the ''intrinsic gene list'', the 70 gene signature, and the CSR signature. By analysis of the intrinsic gene list, it was observed that the majority of basal-like tumours are associated with a poor prognosis signature and a wound response signature, supporting the idea the basal-like tumours are a distinct entity, usually associated with a more aggressive clinical behaviour. When analysed in a multivariate model, only the 70 gene and CSR signatures provided independent and significant prognostic information.[23] The authors also developed a decision tree for coupling the 70 gene and CSR signatures. Firstly, patients were classified according to the former into good or poor prognosis. Secondly, the tumours classified into the poor prognostic group were then classified according to the CSR signature as wound response or quiescent. Those patients with a poor prognosis 70 gene profile but a quiescent CSR signature showed a risk similar to baseline, whereas those patients with both poor prognosis and wound response signature showed a risk of metastatic disease 6.4 fold higher than baseline.[23] This approach shows that combining different signatures with non-overlapping features can be used to strengthen the predictor and may therefore be complementary. However, owing to the limited number of patients in each of the studies performed, interpretation of the results may be overoptimistic.

In a recent study, Wang et al[32] described a new signature, developed for the same purpose of that designed by Van't Veer et al.[28][29] Using an oligonucleotide chip containing 18 400 transcripts (14 500 well characterised human genes), they analysed a series of 286 patients who did not receive systemic therapy; 80 and 206 were randomly assigned to the training and testing sets.[32] The same approach was used to analyse subgroups of oestrogen receptor positive and negative tumours. After developing signatures for each group separately, a final signature composed of 76 genes (60 for the oestrogen positive group and 16 for the oestrogen negative group) was created. This signature showed a specificity of 58% and a sensitivity of 93% for the identification of patients with poor prognosis and proved to be an independent prognostic factor in multivariate analysis for survival without distant metastasis.[32] Like the 70 gene signature, Wang's signature[32] also outperformed the St Gallen[4] and NIH[3] consensus criteria.

Although the potential of the studies summarised above is enormous, it is still unclear whether the final classifier will be composed of a series of signatures or if a there will be a single signature that will outperform the others. As mentioned previously, it is surprising that when comparing different signatures, there is very little overlap between the different gene lists, although some of the differences may be explained by methodological and conceptual differences.[23][32][64] In fact, in the study by Chang et al,[23] it was pointed out that there are no genes concurrently present in the ''intrinsic gene list'', 70 gene signature, and CSR signature. The only overlap observed between these datasets concerns 18 genes present in both the 70 gene signature and ''intrinsic gene list''.[23] In addition, only a three gene overlap between Wang's signature and the 70 gene signature was identified (cyclin E2, origin recognition complex, and tumour necrosis factor superfamily protein).[32]

As pointed out by Jensen and Hovig,[64] devising gene signatures for a given clinical variable is not by itself sufficient to provide significant insight into the underlying biological mechanisms of disease. From the point of view of prognostics, any gene signature with high accuracy and good predictive power would be absolutely fine.[64] However, faced with alternative gene signatures for very similar prediction problems, we are left with the obvious questions of which to trust and why they differ. Given the molecular genetic and histopathological heterogeneity of human breast tumours

and the several questions unanswered, it is still unclear whether internationally recognised gene signatures will be accumulated in the near future.[64]

## MICROARRAYS AS PREDICTIVE FACTORS: SIGNATURES, SIGNATURES EVERYWHERE …

Although the studies above address the issue of which patients may need treatment, a separate question has to be answered: of those who receive treatment, which treatment would be the best choice? Currently, based more on pragmatism rather than scientific evidence, standardised adjuvant chemotherapy is offered. Although hormone receptor status and HER2 status give some direction, systemic treatment is otherwise not tailored to take into account the heterogeneity of tumour biology and response to therapy. Hence, the idea of using expression profiling to identify new predictive signatures and markers of response is quite appealing, and has been applied to tumours treated with a number of different standard neoadjuvant systemic therapy regimens.

A good example of the need of reliable predictive factors is evident in the context of adjuvant tamoxifen therapy.[65] Tamoxifen is the most frequently prescribed antioestrogen agent in women with early stage and metastatic oestrogen receptor positive breast carcinomas. For adjuvant treatment, this is based upon a significant improvement in both recurrence risk and overall survival.[27 65] Tamoxifen given as adjuvant reduces the annual risk of recurrence by 40–50%.[13 27 65] However, almost all patients with metastatic disease and as many as 40% receiving adjuvant tamoxifen eventually relapse due to intrinsic (de novo) or acquired resistance.[65] In a recent study, Ma et al[27] analysed the expression profile of 60 patients uniformly treated with tamoxifen alone using a 22 000 oligonucleotide array. A comparison of response to tamoxifen and expression profiles generated from whole tissue sections and laser capture microdissected samples identified three strongly predictive genes: homeobox gene HOXB13, interleukin 17B receptor (IL17BR) and EST AI240933. The ratio of HOXB13 to IL17BR strongly correlated with recurrence and outperformed other clinical pathological predictors. Validation was then carried out in a cohort of 20 patients using real time, Q-RT-PCR on RNA extracted from paraffin embedded tissue. Furthermore, the authors expressed HOXB13 through a retroviral construct in MCF10A cells and observed that it stimulated cell migration and invasion.[27] Although the number of patients analysed was rather small and requires further validation by independent groups in larger cohorts, it provides a simple ratio of two genes that can be obtained with Q-RT-PCR from formalin fixed, paraffin embedded, tissue sections. It also demonstrates a possible mechanistic role for HOXB13 expression in determining prognosis, illustrating a role of expression profiling studies in candidate gene finding.

In a recent study, Jansen et al[26] compared the expression profiles and response to tamoxifen in a training set of 46 patients. In total, 81genes were differentially expressed in the responsive and non-responsive groups. From these 81 genes, a 44 gene predictor was built and showed significant statistical correlation with response, although it did not reach significance in multivariate analysis. Interestingly, the authors observed an enrichment of genes localised to cytobands 17q21 to 17q22 in the 81 and 44 gene signatures.[26] Although HOXB13 was not in their signatures, this gene maps to 17q21.2. Taken together, these studies suggest that genes localised on the long arm of chromosome 17 may play a role in resistance to tamoxifen; however, the authors failed to assess the impact HER-2 (also positioned on 17q) amplification or immunohistochemical overexpression as a predictor of poor response in their cohort.

The analysis of expression profiling in the neoadjuvant setting is quite appealing.[2 5] Signatures cannot only be devised to predict response to a given chemotherapy agent, but also to compare the profiles before and after treatment, allowing a thorough comparison of the pathways involved. Some studies have used clinical measurements of the tumour as the parameter for response.[66 67] A limitation of this approach is that concordance between clinical and pathological response to cytotoxic therapy may be only moderate. Measurement by palpation can overestimate the number of complete remissions and underestimate the number of non-responders.[68] Futhermore, the dichotomy between clinical partial response and stable disease is arbitrary, and sometimes the cutoff adopted by authors is rather questionable.[69] Even pathological complete response may be limited in its application to defining expression signatures associated with favourable outcome. Although pathological complete response in the neoadjuvant setting is strongly associated with risk of relapse and death of disease, it is still not a perfect surrogate marker for overall survival.[70 71] Perhaps a stricter, uniform definition of pathological complete response is required to reduce subjectivity and strengthen its role in predicting overall survival in the context of both developing new therapies and determining molecular signatures.[70]

Strikingly, several experts have hailed the objectivity of expression profiling when compared with traditional pathology; however, clinical measurement of tumour size has been used in these microarray studies as the means to define response to chemotherapeutic agents. For the reasons outlined above, we wonder how objective and reproducible is this method. Additionally, we have seen anecdotal cases where clinical assessment suggested ''optimum response'' and the pathology specimen showed >90% of viable (non-apoptotic/non-necrotic) tumour cells!

Taxanes, a new class of antimicrotubule agents, have been proven to have equal or greater efficacy than anthracyclines in recent clinical trials.[72 73] Gene expression profile analysis has been used to predict which patients may benefit from taxane therapy. Two classifiers have recently been put forward by two different groups using different methods, namely oligonucleotide microarrays and adaptor tagged competitive PCR.[22 23 25] The overlap between the two classifiers was restricted to three genes; however, the authors have independently claimed that their own signature could predict clinical response to chemotherapy.[22 23 25] These discrepancies may be related to differences in study design and expression profiling methods. Although the results by Chang et al[22 23] and Iwao-Koizumi et al[25] are interesting and may provide further insights in the biology of resistance to docetaxel, it is striking that resistance to docetaxel was linked to the redox system in one study[25] and mTOR survival pathway genes in another.[23] Regrettably, in both studies,[22 23 25] clinical response was used as the surrogate marker and in one of them, the thresholds for response were arbitrarily defined. Caution should be exercised when evaluating the results of studies with a validation set of six patients and a paired comparison of pre-treatment biopsies and surgical samples of only 13 patients.[55 69] Although tempting, scientists should refrain from combining breast carcinomas other than invasive ductal carcinomas under the umbrella descriptive term of ''invasive mammary carcinoma'', as this is not recommended by the World Health Organisation or any other internationally recognised institution.

cDNA microarrays have also been used to develop a signature for response to sequential treatment of paclitaxel followed by fluoracil, doxorubicin, and cyclophosphamide.[53] A 74 gene predictor was devised from a group of 24 patients and then tested in a cohort of 18, which showed an accuracy of 78%, sensitivity of 43%, and specificity of 100%. The results

are very promising and may help us tailor therapy for breast cancer patients.[53] However, the exceedingly small sample size precludes any definite conclusions.[26 55 69] Further validation is eagerly awaited.

Owing to the lack of useful markers to predict response of patients with primary breast cancer to neoadjuvant doxorubicin cyclophosphamide (AC) or doxorubicin docetaxel (AD), a recent study performing expression profiling of tumours in this group of patients was carried out.[34] At variance with previous studies, where fine needle aspiration biopsies were used as the source of RNA and arbitrary definitions of response were adopted, Hannemann *et al*[34] extracted RNA from tissue cores with ≥50% of tumour cells and defined the response to chemotherapy based upon quite well defined pathological and clinical findings. Based upon the expression profile of 46 prechemotherapy samples, neither unsupervised nor supervised methods could separate the responders from nonresponders. When the authors applied hierarchical clustering analysis to prechemotherapy (n = 46) and postchemotherapy (n = 15) samples, they observed that six matched samples clustered together and found that these samples showed either stable disease (n = 5) or partial response (n = 1), suggesting that tumours that responded to chemotherapy showed marked changes in their transcriptomes. Using supervised methods, the authors identified signatures enriched for cell metabolism genes that could differentiate pre and postchemotherapy samples. Although this study could not find a predictive signature for patients treated with AC or AD, the authors adopted better characterised definitions for response and demonstrated that tumours that respond to neoadjuvant chemotherapy show dramatic changes in their expression profiles.[34] Regrettably, the percentage of tumour cells in prechemotherapy and postchemotherapy samples is not described in the report. As the authors defined near pathological complete response as those samples where ''a small number of scattered tumour cells was seen'', it is not clear if different in proportions of stromal cells and tumour cells could have accounted for the distinct expression profiles observed in 7 of 15 matched prechemotherapy and postchemotherapy samples in the responders group.

Although the results outlined above are quite promising, further optimisation of the technique and proper patient selection are needed, before these signatures can be applied to patient management. In fact, a consensus of how clinical trials addressing the impact of these novel high throughput molecular signatures should be designed is awaited.

## A WORD OF CAUTION (OR HOW I LEARNT TO STOP WORRYING AND LOVE MICROARRAYS)

Studies using microarrays as prognostic factors for cancer patients have received a disproportionate attention by leading journals, and the interpretation of the results has been overoptimistic to say the least.[35] Some have dared to say that microarrays will provide data of such quantity and quality that effective treatments or cures for every human disease will be available by 2050.[74] Furthermore, microarray experts have praised clinical trials with a handful of patients and high density data, as opposed to properly designed clinical trials with sufficient number of patients.[75 76] Perhaps, for these experts, ''overfitted'' classifiers may not be an issue.[35]

Recent data by Michiels *et al*[38] have called into question the validity of seven extensively publicised expression profiling studies. Using a simple but insightful approach, these authors showed that in five of these seven studies on array expression and cancer prognosis, the signatures perform no better than tossing a coin![35 38] In addition, these authors observed that in the 500 signatures generated for each training set size, the list of 50 genes with the highest

correlation with outcome was very unstable. Using the dataset of van't Veer *et al,*[29] and the same number of patients in the training set, only 14 of the 70 genes from the published signature were present in more than half of the 500 signatures generated by Michiels *et al*.[38] In addition, >10% of the 4948 genes included in the original dataset were included in at least one estimated signature. Thus, perhaps He and Friend[33] are correct and pathology indeed is unsophisticated and subjective, but as it stands, expression profiling analysis for prognostication of cancer patients certainly requires the skills of a fortune teller.

Each time a new technique is made available, scientists rush and say that it will be the end of histopathology and the excitement of expression profiling has not been an exception. Directly quoting R Klausner (the former director of the National Cancer Institute), it is ''naïve'' to think that ''you could go quickly from this new technology to a clinical tool.''[45] Although we cannot rule out that a new technique may replace pathologists in the future, pathology is still of paramount importance for diagnosis and therapeutic decision making. In a way, the current situation is akin to what happened in the 1980s, when immunohistochemistry was thought to be the end of histopathology. In fact, immunohistochemistry has been embraced by the specialty. Arrays will certainly have a definite impact on our practice, and perhaps the role of pathologists will change slightly, but the specialty is not yet at risk.

....................

**Authors' affiliations**
**J S Reis-Filho, C Westbury,** The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK
**J S Reis-Filho,** Institute of Molecular Pathology and Immunology, University of Porto, Portugal
**J-Y Pierga,** Institut Curie, Section Médicale, Paris, France

## REFERENCES

1 **Simpson PT**, Reis-Filho JS, Gale T, et al. Molecular evolution of breast cancer. *J Pathol* 2005;**205**:248–54.
2 **Cleator S**, Ashworth A. Molecular profiling of breast cancer:clinical implications. *Br J Cancer* 2004;**90**:1120–4.
3 **Eifel P**, Axelson JA, Costa J, 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.
4 **Goldhirsch A**, Wood WC, Gelber RD, et al. Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. *J Clin Oncol* 2003;**21**:3357–65.
5 **Chang JC**, Hilsenbeck SG, Fuqua SA. Genomic approaches in the management and treatment of breast cancer. *Br J Cancer* 2005;**92**:618–24.
6 **Jones C**, Ford E, Gillett C, et al. Molecular cytogenetic identification of subgroups of grade III invasive ductal breast carcinomas with different clinical outcomes. *Clin Cancer Res* 2004;**10**:5988–97.
7 **Lakhani SR**, Ashworth A. Microarray and histopathological analysis of tumours: the future and the past? *Nat Rev Cancer* 2001;**1**:151–7.

8 **Cleator S**, Parton M, Dowsett M. The biology of neoadjuvant chemotherapy for breast cancer. *Endocr Relat Cancer* 2002;**9**:183–95.
9 **Coradini D**, Daidone MG. Biomolecular prognostic factors in breast cancer. *Curr Opin Obstet Gynecol* 2004;**16**:49–55.
10 **Heatley MK**. Ki67 protein:the immaculate deception? *Histopathology* 2002;**40**:483.
11 **Marty M**, Cognetti F, Maraninchi D, et al. Efficacy and safety of trastuzumab combined with docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer administered as first-line treatment: results of a randomized phase II trial by the M77001 study group. *J Clin Oncol* 2005.
12 **Early Breast Cancer Trialists' Collaborative Group**. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;**352**:930–42.
13 **Early Breast Cancer Trialists' Collaborative Group**. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;**351**:1451–67.
14 **Wildsmith SE**, Elcock FJ. Microarrays under the microscope. *Mol Pathol* 2001;**54**:8–16.
15 **Maughan NJ**, Lewis FA, Smith V. An introduction to arrays. *J Pathol* 2001;**195**:3–6.
16 **Sorlie T**, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;**98**:10869–74.
17 **Sorlie T**, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003;**100**:8418–23.
18 **Perou CM**, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
19 **Sotiriou C**, Neo SY, McShane LM, 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.
20 **Bertucci F**, Finetti P, Rougemont J, et al. Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. *Cancer Res* 2004;**64**:8558–65.
21 **Chang HY**, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2004;**2**:E7.
22 **Chang JC**, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003;**362**:362–9.
23 **Chang HY**, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci USA*, 2005;**102**:3738–43.
24 **Fuqua SA**, Chang JC, Hilsenbeck SG. Genomic approaches to understanding and treating breast cancer. *Breast Dis* 2004;**19**:35–46.
25 **Iwao-Koizumi K**, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. *J Clin Oncol* 2005;**23**:422–31.
26 **Jansen MP**, Foekens JA, van Staveren IL et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol* 2005;**23**:732–40.
27 **Ma XJ**, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* 2004;**5**:607–16.
28 **van de Vijver MJ**, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;**347**:1999–2009.
29 **van 't Veer LJ**, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;**415**:530–6.
30 **van't Veer LJ**, Paik S, Hayes DF. Gene expression profiling of breast cancer: a new tumor marker. *J Clin Oncol* 2005;**23**:1631–5.
31 **Huang E**, Cheng SH, Dressman H, et al. Gene expression predictors of breast cancer outcomes. *Lancet* 2003;**361**:1590–6.
32 **Wang Y**, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;**365**:671–9.
33 **He YD**, Friend SH. Microarrays—the 21st century divining rod? *Nat Med* 2001;**7**:658–9.
34 **Hannemann J**, Oosterkamp HM, Bosch CA et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005;**23**:3331–42.
35 **Ioannidis JP**. Microarrays and molecular research: noise discovery? *Lancet* 2005;**365**:454–5.
36 **Ahmed AA**, Brenton JD. Microarrays and breast cancer clinical studies: forgetting what we have not yet learnt. *Breast Cancer Res* 2005;**7**:96–9.
37 **Chang JC**, Hilsenbeck SG, Fuqua SAW. The promise of microarrays in the management and treatment of breast cancer. *Breast Cancer Res* 2005;**7**:100–4.
38 **Michiels S**, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;**365**:488–92.
39 **Kuo WP**, Jenssen TK, Butte AJ, et al. Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* 2002;**18**:405–12.
40 **Dobbin KK**, Beer DG, Meyerson M, et al. Interlaboratory comparability study of cancer gene expression analysis using oligonucleotide microarrays. *Clin Cancer Res* 2005;**11**:565–72.
41 **Larkin JE**, Frank BC, Gavras H, et al. Independence and reproducibility across microarray platforms. *Nat Methods* 2005;**2**:337–44.
42 **Irizarry RA**, Warren D, Spencer F, et al. Multiple-laboratory comparison of microarray platforms. *Nat Methods* 2005;**2**:345–50.

43 **Sherlock G**. Of fish and chips. *Nat Methods* 2005;**2**:329–30.
44 **Bammler T**, Beyer RP, Bhattacharya S, et al. Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2005;**2**:351–6.
45 **Marshall E**. Getting the noise out of gene arrays. *Science* 2004;**306**:630–1.
46 **Tan PK**, Downey TJ, Spitznagel EL Jr, et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 2003;**31**:5676–84.
47 **Zhu B**, Ping G, Shinohara Y, et al. Comparison of gene expression measurements from cDNA and 60-mer oligonucleotide microarrays. *Genomics* 2005;**85**:657–65.
48 **Yauk CL**, Berndt ML, Williams A, et al. Comprehensive comparison of six microarray technologies. *Nucleic Acids Res* 2004;**32**:e124.
49 **Dash A**, Maine IP, Varambally S, et al. Changes in differential gene expression because of warm ischemia time of radical prostatectomy specimens. *Am J Pathol* 2002;**161**:1743–8.
50 **Mutter GL**, Zahrieh D, Liu C, et al. Comparison of frozen and RNALater solid tissue storage methods for use in RNA expression microarrays. *BMC Genomics* 2004;**5**:88.
51 **Auer H**, Lyianarachchi S, Newsom D, et al. Chipping away at the chip bias: RNA degradation in microarray analysis. *Nat Genet* 2003;**35**:292–3.
52 **Imbeaud S**, Graudens E, Boulanger V, et al. Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res* 2005;**33**:e56.
53 **Ayers M**, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004;**22**:2284–93.
54 **West M**, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci USA*, 2001;**98**:11462–7.
55 **Simon R**, Radmacher MD, Dobbin K, et al. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;**95**:14–8.
56 **Korkola JE**, DeVries S, Fridlyand J, et al. Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. *Cancer Res* 2003;**63**:7167–75.
57 **Chang JC**, Wooten EC, Tsimelzon A, et al. Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. *J Clin Oncol* 2005;**23**:1169–77.
58 **Nielsen TO**, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;**10**:5367–74.
59 **Abd El-Rehim DM**, Ball G, Pinder SE, et al. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer* 2005.
60 **Goldhirsch A**, Glick JH, Gelber RD, et al. Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. *J Natl Cancer Inst* 1998;**90**:1601–8.
61 **Tuma RS**. Trial and error: prognostic gene signature study design altered. *J Natl Cancer Inst* 2005;**97**:331–3.
62 **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**:980–91.
63 **Ramaswamy S**, Perou CM. DNA microarrays in breast cancer:the promise of personalised medicine. *Lancet* 2003;**361**:1576–7.
64 **Jenssen TK**, Hovig E. Gene-expression profiling in breast cancer. *Lancet* 2005;**365**:634–5.
65 **Ring A**, Dowsett M. Mechanisms of tamoxifen resistance. *Endocr Relat Cancer* 2004;**11**:643–58.
66 **Cleator SJ**, Makris A, Ashley SE, et al. Good clinical response of breast cancers to neoadjuvant chemoendocrine therapy is associated with improved overall survival. *Ann Oncol* 2005;**16**:267–72.
67 **Pierga JY**, Mouret E, Laurence V, et al. Prognostic factors for survival after neoadjuvant chemotherapy in operable breast cancer: the role of clinical response. *Eur J Cancer* 2003;**39**:1089–96.
68 **Faneyte IF**, Schrama JG, Peterse JL, et al. Breast cancer response to neoadjuvant chemotherapy :predictive markers and relation with outcome. *Br J Cancer* 2003;**88**:406–12.
69 **Brenton JD**, Caldas C. Predictive cancer genomics—what do we need? *Lancet* 2003;**362**:340–1.
70 **Fisher B**, Bryant J, Wolmark N, et al. Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. *J Clin Oncol* 1998;**16**:2672–85.
71 **Ellis M**, Ballman K. Trawling for genes that predict response to breast cancer adjuvant therapy. *J Clin Oncol* 2004;**22**:2267–9.
72 **Nowak AK**, Wilcken NR, Stockler MR, et al. Systematic review of taxane-containing versus non-taxane-containing regimens for adjuvant and neoadjuvant treatment of early breast cancer. *Lancet Oncol* 2004;**5**:372–80.
73 **Estevez LG**, Gradishar WJ. Evidence-based use of neoadjuvant taxane in operable and inoperable breast cancer. *Clin Cancer Res* 2004;**10**:3249–61.
74 **Schena M**. Introduction to microarray analysis. In: Schena M, eds. *Microarray analysis*. New York, NY: Wiley-Liss, 2003:1–26.
75 **Liu ET**. Mechanism-derived gene expression signatures and predictive biomarkers in clinical oncology. *Proc Natl Acad Sci USA*, 2005;**102**:3531–2.
76 **Liu ET**, Karuturi KR. Microarrays and clinical investigations. *N Engl J Med* 2004;**350**:1595–7.