Genomic analyses to select patients for adjuvant chemotherapy: trials and tribulations

B. Weigelt, J. S. Reis-Filho, C. Swanton

Introduction

Breast cancer, the most prevalent malignancy in women, was historically perceived as a single disease, whose clinical behaviour and response to systemic therapy were variable. Treatment decisions were solely based on prognostic clinicopathological parameters, such as tumour size, presence of lymph node metastasis and histological grade, and on three predictive markers of response to endocrine therapy and anti-HER2 agents, namely oestrogen (ER) and progesterone (PR) receptors and HER2, respectively. Importantly, however, predictors of response to adjuvant chemotherapy, the mainstay treatment of up to 60% of patients with breast cancer, have only begun to be incorporated into the armamentarium of clinical decision-making. In fact, only 2%–15% [4] of patients who receive adjuvant chemotherapy based on the clinicopathological characteristics of their tumours benefit from it [1–3].

The selection of patients who should receive chemotherapy is still largely based on patient prognosis (i.e. patients with aggressive disease or a high risk of relapse are offered adjuvant chemotherapy). Although tumour size and lymph node status do not predict the benefit from chemotherapy per se, they are strong prognostic factors; hence, patients with large and/or lymph node-positive tumours still receive chemotherapy based on their baseline risk of relapse. It should be noted, however, that these parameters themselves do not predict the likelihood of the patient to benefit from adjuvant chemotherapy.

Algorithms combining clinicopathological parameters and immunohistochemical markers (e.g. Adjuvant! Online) have proven effective to predict the outcome of cohorts of breast cancer patients and to provide estimates of the benefit that a given patient would derive from endocrine and/or chemotherapy [5, 6]. This approach, however, has proved not to be sufficient for therapy to be tailored to individual patients and substantial disagreement is observed in comparisons among the predictions obtained with Adjuvant! Online and the 21-gene assay Oncotype DX™ [7].

Genomic risk stratification

The advent of high-throughput molecular profiling technologies has led to optimistic promises with regard to the development of molecular tests that can predict important biological and clinical phenomena [8]. The ability to survey the expression levels and copy number of genes in a genome-wide fashion has provided many avenues for the characterisation of the molecular heterogeneity of breast cancers [9–14] and led to the development of multi-parameter tests that are reported to forecast the outcome of patients with breast cancer [15–20].

The ‘paradigm shifting’ contribution of gene expression profiling was the demonstration that ER-positive and ER-negative breast cancers are fundamentally different diseases [21], with distinct risk factors, clinical presentation, natural history and therapeutic response profiles [1–3]. In addition, gene expression profiling has brought to the forefront of cancer research the notion that within the ER-positive and ER-negative groups, there is a great deal of molecular heterogeneity [9–14]. Seminal studies published in the last decade have put forward the existence of the so-called...
‘intrinsic’ subtypes of breast cancer (i.e. basal-like, HER2-enriched, normal breast-like, luminal A and luminal B) [9–13] and other subtypes more recently identified (e.g. molecular apocrine [22, 23] and claudin-low [14]). These subtypes of breast cancer largely correlate with the expression of ER, ER-regulated genes, HER2 and proliferation-related genes [1, 24, 25]. While there is burgeoning evidence for the existence of discrete subgroups of ER-negative disease based on their transcriptomic profiles, the subtypes of ER-positive disease (i.e. luminal A and luminal B) appear to constitute a continuum, whereby luminal A tumours display low levels of proliferation-related genes and high levels of ER-related genes and luminal B cancers harbour the reverse profile [1, 24, 26, 27].

In parallel with the discovery of the molecular heterogeneity of breast cancers, numerous research endeavours have aimed to devise multi-gene predictors or gene signatures (i.e. multi-variable predictors based on the expression levels of specific gene sets) to predict the outcome of breast cancer patients ([15–20] (for reviews see [1, 2, 28]). There is evidence to demonstrate that first generation prognostic signatures, such as Oncotype DX [15], Mammaprint [17, 18], Genomic Grade Index [16] and EndoPredict [20] (Table 1), provide substantial prognostic information for patients with ER-positive breast cancers. This information has been shown to be independent of that provided by conventional clinicopathological parameters (e.g. tumour size, lymph node status and histological grade) [1, 2, 15–17, 19, 20, 28]. Furthermore, some of these predictors have been shown to identify a population of ER-positive breast cancer patients who are unlikely to derive benefit from conventional chemotherapy regimens [29–31]. Recent studies have demonstrated, however, that their prognostic power and their ability to define groups that are unlikely to benefit from chemotherapy largely stems from the expression levels of proliferation-related genes [1, 24, 25]. Importantly, these predictors, however, were developed without taking into account the fact that ER-positive and ER-negative breast cancers are fundamentally different in terms of their transcriptomes and that the levels of expression of proliferation-related genes are only prognostic in ER-positive disease; hence, their usefulness in the prediction of outcome of patients with ER-negative disease is minimal, if any (see below) [1, 24, 32].

**challenges of predicting benefit from adjuvant cytotoxic regimens**

Unlike prognostication, the question of prediction of response or benefit from specific multi-drug adjuvant chemotherapy regimens has proved to be more challenging to answer with empirical and biology-driven studies based on high-throughput technologies [27, 33, 34]. In fact, even within the group of patients ascribed by first generation prognostic signatures as of poor prognosis and potential benefit from chemotherapy, only a proportion benefit from adjuvant cytotoxic regimens [29–31]. Furthermore, drug-specific signatures based on empiricism or from hypothesis-driven *in vitro* studies have proved not to be sufficiently accurate to guide therapy decision-making [33, 34].

Candidate marker approaches have demonstrated that ER and proliferation levels, as defined by immunohistochemical assessment of Ki67, are associated with response to neoadjuvant chemotherapy and correlate with the outcome of breast cancer patients treated with conventional multi-drug chemotherapy regimens [35]. Furthermore, there is burgeoning evidence to suggest that copy number aberrations (i.e. amplification or deletions) of the Topoisomerase IIα gene (TOP2A) may constitute a positive predictor of response to anthracycline-based chemotherapy [36]; however, not all patients who respond to these drugs harbour TOP2A gene amplification [37].

The testing of first generation prognostic signatures as predictive markers of benefit from conventional chemotherapy regimens has demonstrated that the assessment of proliferation-related genes and, to a lesser extent, immune response-related genes does provide potentially clinically useful information for the management of patients with ER-positive breast cancer. The subset of low-proliferation ER-positive breast cancers derives minimal, if any, benefit from chemotherapy [29–31, 38]. In fact, patients classified as of good prognosis by first generation prognostic signatures appear not to derive any benefit from chemotherapy, whereas the subgroup of breast cancer patients who benefit from chemotherapy is almost restricted to the patients classified as of poor prognosis by first generation prognostic signatures [1, 27, 29–31, 38]. Within the group of poor prognosis patients, only a subset benefits from chemotherapy. Hence, this approach offers a high negative predictive value, but a low positive predictive value. Furthermore, the predictive information of these tests above and beyond that offered by standardised Ki67 assessment remains to be fully investigated [8].

The ‘intrinsic’ subtypes of breast cancer, identified by means of hierarchical clustering analysis, microarray-based single-sample predictors, PAM50 and immunohistochemical surrogates, have been shown to correlate with outcome [10–13]. In addition, some of the subtypes have been shown to be associated with substantially higher rates of pathological complete response [39–41] in neo-adjuvant studies where pathological complete response was used as a surrogate for benefit, and with outcome in re-analyses of prospective adjuvant trials [42]. Despite the modest concordance among different methods for the identification of the ‘intrinsic’ subtypes [26, 38, 43–45], some associations have been demonstrated in multiple studies: that up to 50% of basal-like, ‘core basal’ and triple-negative phenotype breast cancers derive substantial benefit from chemotherapy, whereas patients with luminal A tumours derive negligible benefit [39–42]. HER2-enriched breast cancers, which comprise most but not all HER2 amplified breast cancers and do not necessarily harbour HER2 gene amplification [13, 44–46], have also been shown to derive benefit from chemotherapy, in particular from anthracycline-based therapies [42].

As for predictors of response to specific chemotherapy agents derived from hypothesis-driven studies [47] or empirical analyses of tumours, success has been more limited so far. The most promising markers of response to specific chemotherapy agents are TOP2A gene assessment by *in situ*
<table>
<thead>
<tr>
<th>Gene signature</th>
<th>PAM50 (ARUP Laboratories)</th>
<th>Mammaprint™ (Agendia)</th>
<th>Oncotype DX™ (Genomic Health)</th>
<th>Endo-Predict (EP) (Sividon Diagnostics GmbH)</th>
<th>Theros™ (Biotheranostics)</th>
<th>MapQuant DX™ (Ipsogen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay</td>
<td>55-gene PAM50</td>
<td>70-gene signature</td>
<td>21-gene recurrence score</td>
<td>11-gene EP score</td>
<td>2-gene HOBX13:IL17R ratio/5-gene molecular grade index combined into a breast cancer index</td>
<td>97-gene genomic grade index</td>
</tr>
<tr>
<td>Assay platform</td>
<td>qRT-PCR</td>
<td>Microarray</td>
<td>qRT-PCR</td>
<td>qRT-PCR</td>
<td>qRT-PCR</td>
<td>Microarray</td>
</tr>
<tr>
<td>Tissue type</td>
<td>Frozen or FFPE</td>
<td>Frozen or stabilised RNA</td>
<td>FFPE</td>
<td>FFPE</td>
<td>FFPE</td>
<td>Frozen or FFPE</td>
</tr>
<tr>
<td>Results</td>
<td>Classification into the ‘intrinsic’ subtypes: luminal A, luminal B, HER2-enriched and basal-like</td>
<td>Good versus poor prognosis</td>
<td>Low risk, intermediate risk and high risk</td>
<td>Low risk and high risk</td>
<td>Low risk, intermediate risk and high risk</td>
<td>GG1 (low grade); GG3 (high grade)</td>
</tr>
<tr>
<td>Clinical Indication</td>
<td>Prognosis of breast cancer patients, both systemic therapy naive and treated. Prediction of response to anthracycline-based chemotherapy versus CMF</td>
<td>Prognosis in N0, &lt;5 cm diameter, stage I/II, ER+ or ER− BC</td>
<td>Prognosis and prediction of benefit from chemotherapy in ER+ N0/1–3N+ BC on tamoxifen use</td>
<td>ER+/HER2-negative patients treated with endocrine therapy alone</td>
<td>Prognosis and prediction of response to tamoxifen in ER+ BC</td>
<td>Molecular grading in ER+, histological grade II BC</td>
</tr>
<tr>
<td>Level of evidence*</td>
<td>II</td>
<td>III</td>
<td>I</td>
<td>I</td>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>FDA approval</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ASCO recommended</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Randomised trial</td>
<td>No</td>
<td>MINDACT</td>
<td>TAILORx</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

*Levels of evidence as defined by Simon et al. [103].

ASCO, American Society of Clinical Oncology; BC, breast cancer; CMF, cyclophosphamide, methotrexate, fluorouracil; ER, oestrogen receptor; FDA, Food and Drug Administration; FFPE, formalin-fixed paraffin-embedded; N0, lymph node negative; N+, positive lymph nodes; qRT-PCR, quantitative reverse-transcriptase PCR; GG, genomic grade.
hybridisation methods [36] and multi-gene predictors that identify groups of tumours enriched for tumours harbouring TOP2A gene amplification (e.g. the HER2-enriched subtype as identified by PAM50 [42], the subsets of breast cancer defined by the A score [48]).

Many predictive signatures based on empirical analysis of tumours have been developed (for reviews see [1, 2, 27, 28]); however, none of these multi-gene predictors is currently supported by level I evidence or is commercially available. Studies aiming to develop multi-gene predictors of response to multi-drug regimens containing anthracyclines and taxanes have led to the development of signatures that do correlate with response to these drugs [48–54]. These signatures were shown in the training sets (i.e. the microarray experiments used to develop the signature) to predict rather an accurate response to specific combinations of chemotherapy agents, however a reduction in the accuracy of the predictions made by the same multi-gene predictors in independent validation datasets has been observed in the vast majority of cases. Furthermore, independent analyses have demonstrated that their positive and negative predictive values are not sufficient for their introduction in clinical practice [34].

Hypothesis driven approaches, primarily based on the analysis of cancer cell lines (e.g. the NCI60 panel), initially appeared as a fruitful way to identify predictors of response to specific chemotherapy agents [55]. Predictive signatures for a wide gamut of therapeutic agents were developed and applied to the material from a randomised prospective clinical trial [56], which led to the initiation of National Cancer Institute-sponsored prospective clinical trials for their prospective validation. Regrettably, serious doubts about the validity of those observations emerged from the re-analysis of the data by Baggerly and Coombes [33, 57] and independent investigators were unable to confirm the robustness and accuracy of the predictions in independent cohorts [58]. Subsequent investigation led to the retraction of several studies [59] and the termination of the clinical trial that used these predictors.

More elaborate approaches based on the combination of multiple predictors, such as the A score for the prediction of response to anthracyclines [48] and the combined predictor of response to anthracyclines–taxanes [50], are promising; however, additional validation is required before these multi-gene predictors can be further developed into clinical tests.

**limitations of predictive models**

The reasons for the apparent limitations to inability to identify robust predictive signatures are many fold (reviewed in [27, 33]). It is not entirely surprising that while several signatures predictive of response or resistance to specific chemotherapeutic agents indeed have some predictive power in validation studies, their accuracy is not sufficient to define patients who should or should not receive a given agent [27, 33, 60, 61].

From a conceptual standpoint, resistance and sensitivity to a given chemotherapy agent may be caused by a panoply of genetic and/or epigenetic events; hence, these biological phenomena may constitute on their own convergent phenotypes [33, 62–64]. It is, therefore, plausible that a prototypical transcriptomic profile that epitomises resistance to a given therapeutic agent simply does not exist and that, in fact, there might be a variable number of prototypic resistance transcriptomic profiles depending on the number of resistance mechanisms to a given drug. In this case, empirical studies leading to the development of a given gene signature will only be able to identify the most prevalent mechanism of resistance or the mechanism of resistance that results in the most pronounced transcriptomic changes. Hence, searching for a common denominator by microarray analysis may not yield classifiers that are clinically applicable to all patients [27].

Another point to be considered is the observation that resistance to specific chemotherapeutic agents is often caused by alteration of one or a few genes, whose functional changes may not correlate with changes in gene expression profiles, or through post-translational modifications that do not necessarily result in transcriptional differences detected by microarray-based gene expression profiling [27, 33]. Furthermore, resistance to a given therapy may be driven by distinct mechanisms in different molecular subtypes of breast cancer [27, 34]; given that, currently, the actual number of breast cancer molecular subtypes has yet to be defined, it is likely that the signatures generated will not have sufficient accuracy.

A recent study based on ‘spike-in’ experiments of microarray datasets (i.e. a series of artificial gene signatures predicting specific phenomena were artificially introduced in the datasets) demonstrated that there may not be sufficiently highly informative genes to develop clinically useful genomic classifiers to predict certain types of clinical outcomes, including response to chemotherapy agents [65]. Intuitively, sample size would have a strong impact on the ability of identifying gene signatures [65, 66]; although this has been shown to be the case for some biological and clinical phenomena, it has become clear that the number of genes with altered expression and the magnitude of differences in expression levels are more important than the sample size per se [65]. For the identification of robust predictive signatures, numerous genes with strong differences in the expression between sensitive and resistant cases would be required [65].

**inter-tumour heterogeneity**

Next generation sequencing analyses have revealed remarkable inter-tumour heterogeneity in breast cancer. Contrary to copy number aberrations, which have been shown to correlate with ER status and histological grade [67–69], the repertoire of mutations appear to be more variegated and very few somatic mutations are shared among tumours of the same histopathological or clinical subtype [69–73]. In addition, even tumours arising in the context of patients harbouring BRCA1 germline mutations differ in terms of the constellation of somatic genetic aberrations they harbour [71, 74]. Several fusion genes were identified by Stephens et al. [71], none of which were shared in 24 cancers analysed. Inter-tumour heterogeneity, where very few somatic mutations are shared among tumours of the same histopathologic subtype, combined with the polygenic nature of drug resistance mechanisms discussed above, support the contention that drug
resistance mechanisms may be distinct from patient to patient and that the implementation of clinically robust genomic signatures predictive of response to therapies based on the analysis of pre-therapy samples may prove intractable.

intra-tumour heterogeneity: a Darwinian perspective

Current attempts to personalise cancer treatment rely on the accurate genetic profiling from single-tumour biopsies of individual tumours to predict outcome. Intra-tumour genetic heterogeneity (ITH), where individual cancer cells or subpopulations harbouring distinct and divergent genetic or epigenetic abnormalities, may prove an important challenge for biomarker development efforts and the prediction of drug response in advance of therapy [73]. Massively parallel sequencing studies have demonstrated that cancers can be composed of mosaics of tumour cell clones that contain private genetic aberrations in addition to the founder genetic events [75]. Furthermore, several lines of evidence demonstrate that some cancers evolve following a Darwinian evolution model [76–81], with the emergence of multiple subclones that may arise due to an elevated mutation rate followed by selection of the fittest clones at a given time point, based on tumour-specific (e.g. competition and mutualistic relationships between distinct clones), microenvironment-specific (e.g. tumour infiltrating lymphocytes, macrophages and blood supply) and external (e.g. paracrine and endocrine stimuli, therapeutic agents) selective pressures [75, 82]. The tumour and microenvironment interactions may play a role in defining ITH, given the recent evidence suggesting that breast, pancreatic and renal cancers harbour distinct tumour subpopulations that are regionally separated within an individual primary or between primary tumours and metastatic deposits [72, 79, 80, 83, 84]. These data question whether one tumour biopsy from a primary or metastatic site will accurately reflect the genomic landscape of the tumour and raise the question of tumour sampling bias in biomarker studies.

Such concerns undermine the applicability of microarray-based methods to the identification of signatures of resistance or sensitivity to chemotherapy agents due to the composition of tumours of mosaics of neoplastic cells harbouring distinct genetic and epigenetic aberrations that may be regionally separated, in addition to the founder genetic/epigenetic events shared by all cancer cells [72]. In fact, there is direct evidence to demonstrate that at least in some instances, resistance may be driven by selection of a rare subclone from a tumour [33]. Furthermore, emerging evidence suggests that one pattern of ITH, chromosomal instability, may confer intrinsic multi-drug resistance [85, 86], and be associated with distinct patterns of clinical outcome in breast cancer [87, 88]. ITH and its associated impact on drug treatment failure is perhaps best recognised in non-small cell lung cancer (NSCLC) where low-frequency subclonal tumour populations present in the tumour before treatment, harbouring somatic events that initiate drug resistance, are associated with acquired resistance to epidermal growth factor receptor (EGFR) family inhibitors [89, 90]. Such observations support the contention that rare, submodal populations of tumour cells are likely to determine the treatment outcome and conceivably the pattern of metastatic relapse. Given that microarrays provide an average expression profile of a given cancer, this approach is not suitable for the identification of these rare resistant cells.

Although the importance of ITH has only recently been brought to the forefront of breast cancer research, breast cancer ITH has been documented by pathologists for several decades. Up to 15%–20% of breast cancers are classified as of mixed type (i.e. are composed of tumour cells arranged in distinct patterns, where each of the components comprises >10% of the tumour area) [3]. It should be noted, however, that more focal phenotypic heterogeneity is not an uncommon phenomenon in breast cancer [91]. For instance, apocrine areas within an otherwise ductal carcinoma have been shown to harbour distinct patterns of DNA copy number aberrations [92]. Furthermore, metastatic breast cancers, a special histological type that accounts for up to 5% of all invasive breast cancers, are characterised by a complex admixture of invasive components exhibiting histological patterns other than glandular differentiation. In metastatic breast carcinomas, morphologically distinct components may harbour distinct patterns of gene copy number aberrations. We have recently demonstrated that, in a subset of metastatic carcinomas, certain genetic aberrations segregate with specific morphological components of a given tumour, providing direct evidence of ITH in these cancers and suggesting that specific genetic aberrations may either underpin or at least be coincidental with distinct histological patterns [84].

Characterising ITH is likely to have an important impact on our understanding of processes that regulate disease progression, resistance to specific therapeutic agents and treatment failure. For example, ITH of microsatellite marker allelic loss appeared to be relatively higher in ductal carcinoma in situ (DCIS) associated with invasive lesions than those in pure DCIS samples, suggesting that ITH may play a role in the genetic divergence required for evolutionary robustness during disease progression [93]. Furthermore, in synchronous DCIS and invasive carcinomas, there is evidence to suggest that the invasive component emerges from a genetically distinct subpopulation of the DCIS [94].

ITH that appears to occur in spatially separated regions of the same tumour both at the macroscopic and genetic level question the reliability of single-tumour biopsies for genomic profiling to predict outcome. Through the analysis of chromosome 22q DNA copy number aberrations in multiple biopsies from the same breast tumours, Benetkiewicz et al. provided evidence questioning single-biopsy genomic studies. In 15 tumours with 2 to 3 biopsies per tumour, 26.6% of tumours harboured intratumour heterogeneity for 22q aberrations [95]. Similarly, Torres et al. [96], through sampling different regions of the same tumours, provide evidence that ITH in both the number and type of genomic imbalances, is common in breast cancer. Disturbingly, ITH has been substantiated in breast cancer, not only between spatially separated regions of the same tumour but also within single-breast cancer biopsies; through the use of single-cell analytical techniques, multiple tumour subpopulations occur within...
single biopsies, each with distinct DNA copy number profiles [97, 98].

The degree of ITH is likely to be cancer type-specific and is probably subtype-specific, in the context of breast cancer [82]. Genetic diversity in breast cancer is likely to be driven by the type and level of genetic instability a given tumour may harbour. For instance, ER-negative breast cancers not uncommonly display a ‘mutator’ phenotype characterised by multiple tandem duplications, whereas this phenomenon israrer in ER-positive cancers [71]. In addition, loss of function of genes that control certain types of DNA repair appear to be more prevalent in ER-negative disease, and may constitute one of the causes for the high levels of genetic instability in this subtype of breast cancers [82].

The fact that the multi-gene predictors have so far been derived from the analysis of pre-therapy primary tumour samples and are meant to predict the behaviour of disseminated cells may be one of the reasons for the limited success in this field. If the genetic make-up of a tumour evolves following a Darwinian model, the tumour cell population obtained from the primary tumour bulk and analysed by gene expression profiling may not be representative of the metastatic clone [75, 82]. Consistent with this notion, two massively parallel sequencing-based studies have recently reported important differences in the repertoire of somatic genetic aberrations between primary breast cancers and their metastases [70, 72] with similar results in other tumour types [80, 99]. Furthermore, analysis of individual markers, such as ER, PR and HER2, have demonstrated that in a substantial minority of cases, the profile of the relapse may significantly differ from that of the primary tumour [100, 101].

other models of drug resistance

Conclusions based on the relevance of ITH, poor outcome and drug resistance should be guarded. It is clear that drug resistance and poor outcome do not exclusively follow a Darwinian model driving the selection of low-frequency drug-resistant clones. For example in ER-positive breast cancer, through the mapping of ER-binding events by ChIP-seq, drug resistant cancers are typified by unique ER-binding regions in poor-risk tumours, driven by FOXA1, rather than the selection of a low-frequency subpopulation of resistant cells [102]. Conceivably, such general phenomena, rather than low-frequency events driving poor outcome, in ER-positive breast cancer explains some of the limited successes the field has witnessed in the genomics biomarker field.

conclusions

Rare tumour subpopulations that may be selected for during drug therapy or metastatic progression may be relatively hidden and evade detection within the tumour through summary genomic measures acquired through microarray-based expression profiling or DNA copy number analyses that only surveys modal populations. Characterising ITH using techniques that can begin to identify rarer tumour subpopulations and the change in the subclonal architecture of the tumour through therapy is likely to have important consequences on our understanding of the processes that regulate the acquisition of drug resistance and treatment failure [81]. Development of clinically implementable techniques to identify such low-frequency events that may influence clinical outcome in solid tumours is likely to provide an exciting research area for progress in this field.

disclosure

The authors have declared no conflicts of interest.

references


75. Swanton C, Caldas C. From genomic landscapes to personalized cancer management—is there a roadmap? Ann N Y Acad Sci 2010; 1210: 34–44.


