Current and emerging biomarkers in breast cancer: prognosis and prediction

Marion T Weigel¹ and Mitch Dowsett¹,²

¹The Institute of Cancer Research, Breakthrough Breast Cancer Research Centre, 237 Fulham Road, London SW3 6JB, UK
²Royal Marsden Hospital, Fulham Road, London SW3 6JJ, UK
(Correspondence should be addressed to M Dowsett; Email: mitch.dowsett@icr.ac.uk)

Abstract

Breast cancer treatment has experienced several changes in the past decades due to the discovery of specific prognostic and predictive biomarkers that enable the application of more individualized therapies to different molecular subgroups. These subgroups show specific differences regarding biological clinical behavior. In addition to the classical clinical prognostic factors of breast cancer, established molecular biomarkers such as estrogen receptor and progesterone receptor have played a significant role in the selection of patients benefiting from endocrine therapy for many years. More recently, the human epidermal growth factor receptor 2 (HER2) has been validated to be not only a prognostic factor, but also a predictor of response to HER2 targeting therapy. The shift toward an earlier diagnosis of breast cancer due to improved imaging methods and screening programs highlights the need for new factors and combinations of biomarkers to quantify the residual risk of patients and to indicate the potential value of additional treatment strategies. The marker of proliferation Ki67 has recently emerged as an important marker due to several applications in neoadjuvant therapy in addition to its moderate prognostic value. With the introduction of high-throughput technologies, numerous multigene signatures have been identified that aim to outperform traditional markers: current prospective clinical trials are seeking evidence for their definitive role in breast cancer. There exist many more factors and approaches that have the potential to become relevant in the near future including the detection of single disseminating and circulating tumor cells in blood and bone marrow as well as of circulating cell-free DNA and microRNA. Careful randomized prospective testing and comparison with existing established factors will be required to select those emerging markers that offer substantial cost-effective benefit and thereby justify their routine use for breast cancer therapy decision-making.

Introduction

Breast cancer is still one of the leading causes of cancer death in women, but there has been a sustained decline in mortality rates over the last decades. The incremental application of increasingly effective adjuvant medical treatments is one of the major factors for this development, despite an increasing incidence of breast cancer. As a result of regular mammography screening programs, a shift toward the detection of early-stage (<2 cm) node-negative breast cancer with better prognosis has occurred. While this almost certainly also contributes to improved outcomes, it also poses a challenge for clinicians regarding the choice of optimal adjuvant treatment. The relapse rate after surgery alone in patients detected with early breast cancer is relatively low, and the individual estimation of the absolute benefit of systemic chemotherapy has to be taken in consideration when making therapeutic decisions. It is of great importance to avoid overtreatment in patients who only receive a modest benefit, while suffering from toxic side effects. On the other hand, undertreatment or incorrect treatment also has to be avoided. It is therefore necessary to define specific characteristics, which provide the possibilities for individual treatment optimization.

Classical clinicopathological features indicating patient prognosis include tumor size, histological...
subtype and grade, lymph node metastases, and lymphovascular invasion, which are derived from careful histological analysis of primary breast cancer samples. The TNM (tumor size, nodes, metastasis) system integrates these into tumor stages that have major prognostic value (Table 1). But in this era of high-throughput methods, a deluge of novel biomarkers have been reported for prognostic and predictive purposes. But out of these, only a few have made their way into clinical routine due to the lack of sufficient validation to reach a Level of Evidence I or II according to the American Society of Clinical Oncology’s Tumor Marker Utility Grading System (Hayes et al. 1996, Harris et al. 2007). Using this system, only two biomarkers, estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2), have been established and are assessed routinely in every breast cancer. Nonetheless, the identification of new markers has led to a more definitive insight into tumor biology and substantiates the importance of the existing biomarkers. To facilitate future research on biomarkers, a guideline named reporting recommendations of tumor marker prognostic studies (REMARK) has been published, which recommends a description of the amount of information that should be provided when reporting the results of biomarker studies (McShane et al. 2005).

In this review, we discuss the importance of established prognostic factors and predictive biomarkers as well as some emerging biomarkers that are currently undergoing testing for technical validity and clinical utility.

### Prognosis and prediction

Prognostic and predictive markers are both of high relevance in therapeutic decision procedures in order to individualize treatment, but they have distinct roles. Prognostic and predictive factors may be derived from either the characteristics of the patient or the tumor

<table>
<thead>
<tr>
<th><strong>Table 1</strong> TNM classification for breast cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T (tumor size)</strong></td>
</tr>
<tr>
<td>TX</td>
</tr>
<tr>
<td>T0</td>
</tr>
<tr>
<td>Tis</td>
</tr>
<tr>
<td>T1</td>
</tr>
<tr>
<td>T1mic: microinvasion 0.1 cm or less in greatest dimension</td>
</tr>
<tr>
<td>T1a: tumor more than 0.1 cm but not more than 0.5 cm in greatest dimension</td>
</tr>
<tr>
<td>T1b: tumor more than 0.5 cm but not more than 1.0 cm in greatest dimension</td>
</tr>
<tr>
<td>T1c: tumor more than 1.0 cm but not more than 2.0 cm in greatest dimension</td>
</tr>
<tr>
<td>T2</td>
</tr>
<tr>
<td>T3</td>
</tr>
<tr>
<td>T4</td>
</tr>
<tr>
<td>T4a: extension to chest wall</td>
</tr>
<tr>
<td>T4b: edema (including peau d’orange) or ulceration of the skin of the breast or satellite skin nodules confined to the same breast</td>
</tr>
<tr>
<td>T4c: both of the above (T4a and T4b)</td>
</tr>
<tr>
<td>T4d: inflammatory carcinoma</td>
</tr>
</tbody>
</table>

| **N (nodes)**                                  |
| NX | Regional lymph nodes cannot be assessed (e.g. previously removed) |
| N0 | No regional lymph node metastasis              |
| N1 | Metastasis to movable ipsilateral axillary lymph node(s) |
| pN1a: only micrometastasis (none larger than 0.2 cm) |
| pN1b: metastasis to lymph node(s), any larger than 0.2 cm |
| pN1bi: metastasis to one to three lymph nodes, any more than 0.2 cm and all <2.0 cm in greatest dimension |
| pN1bii: metastasis to four or more lymph nodes, any more than 0.2 cm and all <2.0 cm in greatest dimension |
| pN1biii: extension of tumor beyond the capsule of a lymph node metastasis <2.0 cm in greatest dimension |
| pN1biv: metastasis to a lymph node 2.0 cm or more in greatest dimension |
| N2 | Metastasis to ipsilateral axillary lymph node(s) fixed to each other or to other structures |
| N3 | Metastasis to ipsilateral internal mammary lymph node(s) |

| **M (distant metastasis)**                     |
| MX | Presence of distant metastasis cannot be assessed |
| M0 | No distant metastasis                           |
| M1 | Distant metastasis present (includes metastasis to ipsilateral supraclavicular lymph nodes) |

p, pathologically determined.
type. Prognostic factors intend to predict objectively and independently patient clinical outcome independent of treatment, while predictive factors aim to foretell the response of a patient to a specific therapeutic intervention and are associated with tumor sensitivity or resistance to that therapy. Prognostic factors necessarily require definition in patient cohorts that did not undergo systemic adjuvant treatment.

Predictive markers may be the target of a specific therapy itself. For example, the oncogene HER2 is the target of the monoclonal antibody trastuzumab, and HER2 amplification predicts for a good response to anti-HER2 therapy. It is important to note that HER2 status is also prognostic, and like many factors, it has mixed prognostic/predictive significance. Similarly Ki67, which will be discussed in more detail below, as a marker of proliferation displays a strong prognostic effect, but it also appears to predict for a good response to systemic chemotherapy. In general, prognostic markers help to determine whether a patient requires treatment, and a predictive factor is useful in deciding which treatment will be the best.

Lately, there has been an increase in implementation of marker combinations to define treatment-specific prognoses. This is of special interest to define the residual risk of recurrence when a patient is treated in a specific manner and to evaluate the potential importance of further treatment options. Great efforts, especially on the transcriptome level, have been made to discriminate which ER-positive early breast cancer patients would really benefit from additional chemotherapy and who could be spared of it and the side effects. The ability to do this was highlighted as the top priority question in breast cancer translational research in an international survey of breast cancer specialists (Dowsett et al. 2007a).

Established biomarkers

Estrogen receptor

ER (α) expression is undoubtedly the most important biomarker in breast cancer, because it provides the index for sensitivity to endocrine treatment. ER-positive tumors (c. 80% of breast cancer) use the steroid hormone estradiol as their main growth stimulus; ER is therefore the direct target of endocrine therapies. The Oxford overview confirms that patients with ER-negative disease have no benefit from 5-year adjuvant treatment with tamoxifen, but some benefit may be derived in the uncommon group of ER-negative and progesterone receptor (PgR)-expressing breast tumors (2005). In contrast, such treatment reduces the annual breast cancer death rate by 31% in ER-positive disease.

While the absence or presence of the ER is used to obtain treatment decisions, little attention has been paid on the value of the quantitative expression levels as a predictive indicator. Evidence from the 1970s suggests a direct correlation between ER expression levels and response to endocrine therapy (Byar et al. 1979). The Early Breast Cancer Trialists’ Collaborative Group reported that higher levels of ER were associated with a lower risk of recurrence when receiving adjuvant tamoxifen (1998). Similar results were obtained in the NSABP-14 trial using the ligand-binding assay and mRNA expression of ER (Paik et al. 2004). More recent analyses from the large prospective adjuvant trials anastrozole, tamoxifen, alone or in combination (ATAC) and BIG 1-98 (letrozole versus tamoxifen) comparing aromatase inhibitors (AIs) with tamoxifen did not find a subgroup of ER-positive patients with different ER expression levels, which derives a greater benefit from AIs versus tamoxifen (Viale et al. 2007, Dowsett et al. 2008). The trials revealed, however, that higher ER levels were related to improved outcome of both the endocrine treatments (Fig. 1).

It has been reported that ER status predicts for response to chemotherapy in the neoadjuvant setting. Multiple clinical studies have demonstrated that the ER-negative breast cancer patients are more likely to achieve a pathological complete response (pCR) with neoadjuvant chemotherapy than the ER-positive patients, with pCR rates of 7–8 vs 21–33% respectively being reported (Colleoni et al. 2004, Ring et al. 2004). This may be partly explicable by the ER-negative breast tumors tending to have higher proliferation rates, but this does not appear to provide a full explanation (Jones et al. 2009).

There have also been investigations concerning the amplification of the ER gene (ESR1). An initial report indicated that ESR1 gene amplification in breast cancer could be detected in ~20% of all invasive tumors, and that there was a correlation between the gene amplification and ER expression levels (Holst et al. 2007). However, <3% of invasive breast cancer cases were reported as ESR1 amplified by other independent groups (Brown et al. 2008, Horlings et al. 2008, Reis-Filho et al. 2008).

Extensive research has been undertaken in trying to discover the function and relevance of different splice variants and point mutations of the ER. One ER mutation (K303R), which leads to a receptor that is able to induce proliferation even in conditions of low
hormone levels, has been reported as being associated with benign breast hyperplasia and breast cancer by one group (Herynk & Fuqua 2004) but not confirmed by others. Despite much investigation of ESR1 mutations and splice variants, their clinical role appears to be small.

The recently released guideline of the American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) has the aim to improve hormone receptor testing for patients with breast cancer, and recommends ER and PgR testing in all newly diagnosed cases as well as in any local or distant recurrence whenever appropriate (Hammond et al. 2010).

**Progesterone receptor**

The expression of the PgR is strongly dependent on the presence of ER. Tumors expressing PgR but not the ER are uncommon and represent <1% of all breast cancer cases in some large series (Viale et al. 2007). For this reason, tumors with PgR expression lacking ER expression should undergo a retesting of their ER status to eliminate false ER negativity. In the rare cases of solely PgR-expressing patients, some limited benefit from tamoxifen is described, but endocrine therapy is still widely recommended (Dowsett et al. 2006a).

There is evidence that in metastatic breast cancer the response to anti-estrogen treatment is better among patients with tumors expressing both ER and PgR versus those who only show ER positivity but lack the PgR expression (Elledge et al. 2000). Data from adjuvant trials comparing tamoxifen treatment with controls indicate a strong prognostic value for PgR expression, but indicate a little predictive significance (Dowsett et al. 2006a). Patients with high levels of PgR within their breast tumors have a better outcome than low expressors with tamoxifen, but the relative benefit from tamoxifen remains similar (2005, Dowsett et al. 2006a).

The impact of PgR expression on response to and outcome of treatment with AIs has been less clear.
The ATAC trialists published a hypothesis generating report suggesting that patients with PgR-negative breast cancer would obtain a substantially greater benefit from anastrozole than from tamoxifen compared with PgR-positive patients (Dowsett et al. 2005a). However, this hypothesis was not confirmed in centrally analyzed material from 1856 ER- and/or PgR-expressing patients (Fig. 1). Moreover, the BIG 1-98 trial reported that the benefit from letrozole over tamoxifen did not vary according to the PgR status (Viale et al. 2007). Nevertheless, these adjuvant trials clearly supported the existence of a strong relationship between PgR expression levels and prognosis on endocrine therapy, which may be useful in estimating residual risk.

HER2

The oncogene HER2 was first identified to be an indicator of patient’s prognosis. In cases of HER2 being overexpressed (HER2 positive), breast cancer patients are more likely to suffer from relapse and tend to have a shorter overall survival. Amplification of the HER2 gene and RNA/protein overexpression correlate strongly (Pegram et al. 2000). Through the development of the monoclonal antibody trastuzumab, which is targeted at HER2, the amplification status of HER2 became also a highly predictive biomarker (Slamon et al. 1987, Mass et al. 2005). Overexpression and amplification of HER2 can be detected in about 15% of all primary breast cancers, and this group of patients benefit significantly from anti-HER2 therapies. HER2 status should be assessed in every diagnosed case of breast cancer (Romond et al. 2005, Smith et al. 2007).

HER2 status is currently assessed in most cases initially by immunohistochemistry, and in cases of equivocal protein expression levels, HER2 gene copy number is measured via fluorescence in situ hybridization (FISH) or chromatin in situ hybridization (CISH) techniques (Wolff et al. 2007). Usually, when using ISH techniques, two locus-specific probes, one for HER2 and the other for the centromere of chromosome 17 (CEP17), are applied, and the ratio of the two is calculated. A ratio > 2.2 is considered unambiguously positive. The need for accurate application of anti-HER2 therapy to the sensitive population highlights the importance of accurate testing. This led to the creation of the American Society of Clinical Oncology/College of American Pathologists guidelines on methodology for immunohistochemistry and ISH techniques for establishing gene copy number of HER2 as well as on test interpretation (Wolff et al. 2007).

HER2 amplification or gain beneath the threshold of two gene copies per CEP17 seems to have only little or no significance in terms of prognosis and prediction of benefit from anti-HER2 treatment in the adjuvant setting (Dowsett et al. 2009), but some uncertainty in the precise level of this threshold has arisen. This is due to a small subgroup included in the NSABP-B31 trial on the basis of local HER2-positive status that showed a significant benefit from trastuzumab, despite central analysis of their tumors revealing HER2 negativity (Paik et al. 2008).

Recent studies also describe an association of HER2 amplification with benefit from adjuvant doxorubicin-based chemotherapy (Muss et al. 1994, Dressler et al. 2005) as well as from paclitaxel administered after four cycles of doxorubicin plus cyclophosphamide (Hayes et al. 2007). The benefit of HER2-positive patients from anthracyclines appears to be exclusive to the subgroup with HER2-amplified tumors, although some data suggest that this may be due to a co-amplification of the topoisomerase IIα gene (Gennari et al. 2008). Yet more recent data suggest that this relationship of HER2 and/or topoisomerase IIZ with anthracycline response may be due to polysomy of chromosome 17 rather than due to amplification of the genes per se (Bartlett et al. 2010).

A poorer response to tamoxifen has been reported among hormone receptor and HER2-positive patients, but the impediment to response is insufficient for the selection of endocrine treatment agents to be influenced by the HER2 status (Ring & Dowsett 2004). Neither the ATAC nor the BIG 1-98 trials observed differences in the benefit between tamoxifen and AIs according to HER2 status (Viale et al. 2007, Dowsett et al. 2008).

Within the last few years, many more promising agents targeting HER2 have been developed including monoclonal antibodies and tyrosine kinase inhibitors. For them, HER2 status is likely to be a predictive marker as well (Widakowich et al. 2008). The preliminary finding that patients expressing a truncated cytoplasmic HER2 receptor (p95HER2) show a poor response to trastuzumab (Anido et al. 2006), but may benefit from the tyrosine kinase inhibitor lapatinib (Scaltriti et al. 2007), might have some implications for future HER2 testing.

Emerging biomarkers

Ki67

The marker of proliferation Ki67 was first identified by Gerdes et al. (1983) in the 1980s using a mouse monoclonal antibody against a nuclear antigen from a
Hodgkin’s lymphoma cell line. Ki67 is a nuclear non-histone protein and was named after the researcher’s location. In this context, Ki stands for the University of Kiel, Germany, and 67 refers to the number of the clone on the 96-well plate. The characteristic that Ki67 was universally expressed among proliferating cells and absent in quiescent cells led to the further evaluation of Ki67 as a marker of proliferation. Although little is known about the exact function of the protein in cell division, Ki67 is expressed during G1, S, and G2 phases of cell cycle with a peak during mitosis and an absence in G0 phase (Lopez et al. 1991).

Although initially the Ki67 antibody was applied only to fresh frozen tissue, many more antibodies with applicability in paraffin-embedded tissue were developed with MIB-1 (targeting the same epitope as the original one) being the most frequently used (Harris et al. 2007). Ki67 expression levels are determined as the percentage of tumor cell nuclei positively stained. Until now, no absolute standard methodology and cut-off point have been defined.

The correlation of Ki67 and other biomarkers in invasive breast cancer has been studied intensively. Unsurprisingly, given that the Nottingham grading system defines mitotic rate as one of its three criteria (Trihia et al. 2003), there is a good correlation with tumor grade. The relationship with ER has been predominantly described as an inverse correlation with lower proliferative activity in ER-positive tumors (Haerslev et al. 1996). There are hints of a correlation with HER2 as well, but this is not completely defined (Nicholson et al. 1993, Rudolph et al. 1999).

A possible prognostic role for proliferation marker Ki67 in breast cancer has been investigated in many studies. A review of 40 studies involving more than 11 000 patients presented by our group (Urruticoechea et al. 2005) describes strong evidence for the ability of Ki67 as a single variable to distinguish between a good or bad outcome in the group of node-negative patients. Unfortunately, this ability is not maintained in the multivariate analyses in all of the included studies. Another meta-analysis including disease-free survival data from 29 studies confirmed the adverse effect on overall survival and disease-free survival in cases of positive staining for Ki67 among both the node-negative and node-positive breast cancer (de Azambuja et al. 2007). Although the most recently published analysis of 15 790 cases from 43 studies reported an association of Ki67 positivity with shorter overall survival, Ki67 staining is still not recommended as a prognostic marker for routine use (Stuart-Harris et al. 2008).

New approaches of combining established markers with novel factors are currently under evaluation. One of these is an immunopanel of ER, PgR, HER2, and Ki67, whose ability to distinguish between luminal A and B subtypes in a similar manner as the original 50-gene signature has been shown (Cheang et al. 2009). In this context, a cut-off of 13.25% Ki67 positive staining was used to discriminate between the subtypes, meaning that a higher score defines luminal B tumors with a worse prognosis.

In early as well as locally advanced breast cancer, baseline Ki67 has been found to predict for response to chemotherapy, whereas this is not the case for endocrine treatment (Chang et al. 2000a,b, Faneyte et al. 2003). Findings of our group indicate that post-neoadjuvant chemotherapy measurement of Ki67 is a strong predictor for recurrence-free and overall survival. However, a high pretreatment score is associated with a good chance to achieve a pCR, and this is a predictor of long-term outcome in these patients (Jones et al. 2009).

In the BIG 1-98 adjuvant trial of letrozole versus tamoxifen, the absolute benefit of the AI over tamoxifen was greatest at the highest levels of Ki67 (Viale et al. 2007). This was due predominantly to the poorer prognosis of those patients with there being a significant increase in relative benefit.

Recent neoadjuvant endocrine studies have evaluated the use of serial Ki67 measurements. These have found that the detection of changes in Ki67 predicts for treatment benefit and highlighted the role of measuring Ki67 early on-treatment as a superior predictor of long-term outcome than pretreatment expression. In this context, the results of the IMPACT trial demonstrated that the decrease in Ki67 was greater at 2 and 12 weeks of AI treatment than treatment with tamoxifen or the combination of the two drugs (Dowsett et al. 2005b,c, 2006b). Thus, the IMPACT trial showed similarly greater efficiency of the AI versus tamoxifen and the combination as observed in the adjuvant ATAC trial, where the anastrozole alone arm showed a prolonged recurrence-free survival (Howell et al. 2005). The adjuvant trial required 30 times as many patients and 10 times the follow-up compared with the neoadjuvant trial to make its first report on efficacy. While on-treatment measurements of Ki67 will not replace the need for adjuvant trials, they do provide an approach to quickly test the potential effectiveness of candidates for phase III trials. Moreover, higher Ki67 levels after 2 weeks of endocrine treatment were linked to shorter recurrence-free survival, and the pretreatment value added no extra prognostic information.
The evaluation of this concept is currently conducted in the perioperative endocrine treatment for individualizing care (POETIC) trial. The main biomarker aim of this trial is to determine whether the measurement of Ki67 after 2 weeks of presurgical treatment with AI is sufficiently more predictive than in the absence of treatment to merit introducing this into routine clinical practice. In the case of a positive result, this has the potential to radically change the assessment of prognostic markers.

**Cyclin D1**

Cyclin D1 is overexpressed at the mRNA and protein level in over 50% of breast cancer cases including 15% in which a gene amplification occurs (Buckley et al. 1993, Gillett et al. 1994, Ormandy et al. 2003). Cells in the G1 phase of the cell cycle react to growth factor stimulation with the induction of D-type cyclins (Musgrove et al. 1993). Subsequently, cyclin D1 binds to cyclin-dependent kinases leading to phosphorylation of various substrates including RB protein (Matsushime et al. 1994). This contributes to the regulation of G1–S phase transition in the cell cycle. In particular, cyclin D1 has the ability to regulate the proliferation of estrogen-responsive cells (Zwijnsen et al. 1997), and a strong positive correlation with ER and PgR expression levels has been described (Hui et al. 1996, Barbareschi et al. 1997, Jares et al. 1997). Importantly, while there is strong evidence that overexpression of cyclin D1 is a prognostic factor for better outcome in invasive breast cancer, in particular among ER-positive patients (Gillett et al. 1996, Hwang et al. 2003, Bilalovic et al. 2005), its amplification is associated with early relapse and poor prognosis (Michalides et al. 1996, Seshadri et al. 1996, Bieche et al. 2002). A possible predictive value of cyclin D1 in hormone receptor-positive patients has also been shown recently: overexpression as well as amplification is a predictor of poor response to anti-estrogen treatments (Stendahl et al. 2004, Jirström et al. 2005). Cyclin D1 merits further investigation in this context.

**Cyclin E**

Cyclin E acts similarly to cyclin D1 as a positive regulator of cell cycle transition with peak levels of protein expression and enzymatic complex formation with cyclin-dependent kinase 2 in the G1 phase (Koff et al. 1992). Cyclin E gene amplification has been detected in several breast cancer cell lines, and there is strong evidence that cyclin E plays a role in tumorigenesis (Buckley et al. 1993, Keyomarsi & Pardee 1993, Bortner & Rosenberg 1997). The full-length protein is altered by post-translational cleavage resulting in hyperactive low molecular weight forms, which are uniquely detectable in tumor cells and correlate with increasing stage and grade of breast cancer (Keyomarsi et al. 1994, Nielsen et al. 1996, Wang et al. 1996, Wingate et al. 2009). The clinical significance of cyclin E has been studied repeatedly. One study measured cyclin E expression in 395 primary breast cancer cases, and correlated the data with established prognostic factors and clinical outcome. Both low molecular weight and total cyclin E levels emerged as the most powerful discriminants of overall and disease-free survival outperforming classical clinical and pathological biomarkers in univariate and multivariate analyses (Keyomarsi et al. 2002). The role of cyclin E in cell cycle suggests that increased levels may alter the response to chemotherapy and endocrine therapy. It has been

(Fig. 2; Dowsett et al. 2007b). The evaluation of this concept is currently conducted in the perioperative endocrine treatment for individualizing care (POETIC) trial. The main biomarker aim of this trial is to determine whether the measurement of Ki67 after 2 weeks of presurgical treatment with AI is sufficiently more predictive than in the absence of treatment to merit introducing this to routine clinical practice. In the case of a positive result, this has the potential to radically change the assessment of prognostic markers.
shown that altered levels increase the sensitivity of breast cancer cells to cisplatin and paclitaxel effects (Smith & Seo 2000), but on the other hand facilitate breast cancer cells to cisplatin and paclitaxel effects shown that altered levels increase the sensitivity of breast cancer cells to cisplatin and paclitaxel effects (Akli & Keyomarsi 2004). These results have to be validated in prospective trials.

**ERβ**

First discovered in 1996, ERβ is frequently expressed in cells of different organs. Although the two ER subtypes are encoded by different genes located on two different chromosomes, they have much in common structurally. The ligand-binding domain in the subtypes exhibits a 59% homology (Enmark et al. 1997, Matthews & Gustafsson 2003). Both receptors tend to bind to estrogen-response elements with similar affinity due to a high homology of their DNA-binding domains. Remarkably, ERβ has the ability to mediate sometimes opposite effects to ERα due to different binding regions (Liu et al. 2008) and the existence of different splice variants of ERβ. Especially, ERβ cx is known to exhibit a dominant-negative activity against ERα (Herynk & Fuqua 2004).

A role of ERβ in tumorigenesis has been suggested due to its significant down-regulation in breast cancer compared with normal breast tissue in contrast to ERα (Roger et al. 2001, Skliris et al. 2003, Bardin et al. 2004). A strong down-regulation of c-myc, a regulator of cell cycle and transcription, can be observed in the presence of ERβ, which may explain ERβ’s anti-proliferative effects (Ström et al. 2004). Initially, ERβ mRNA levels were used to evaluate its prognostic value. This led to contrary results, associating ERβ with poor prognosis and endocrine resistance on one side, and a good prognosis on the other side (Leygue et al. 1998, Speirs et al. 1999, Park et al. 2003). The availability of specific antibodies made it possible to relate ERβ protein levels with good prognosis, prolonged disease-free survival, and response to tamoxifen (Skliris et al. 2003, Hopp et al. 2004). Moreover, ERβ expression has been linked to the expression of ERα and PgR (Järvinen et al. 2000, Omoto et al. 2001), and in ERα -negative breast cancer, ERβ is expressed in about 50% of the cases (Skliris et al. 2006). In this setting, ERβ correlates positively with the marker of proliferation Ki67, and an association with HER2 overexpression has been described (Jensen et al. 2001, Skliris et al. 2006, Umekita et al. 2006). However, further studies are needed before ERβ could be used as a diagnostic tool and possible target of therapy.

**Multigene parameters**

In the last two decades, the human genome project and the development of high-throughput technologies have paved the way for the so-called ‘-omics’ revolution. Gene expression profiling of tumors allows the measurement of thousands of mRNA transcripts in one single experiment using DNA microarrays. The recent St Gallen consensus statement provides the view that the use of a validated multigene profiling assay is warranted as an adjunct to high-quality phenotyping of breast cancer in cases in which the indication for adjuvant chemotherapy remains uncertain (Goldhirsch et al. 2009).

In breast cancer, the results of these expression profiling studies indicated the existence of a number of molecularly distinct neoplastic disorders, which appear to originate from different cell types. Perou et al. (2000) were the first to distinguish four molecular classes of breast cancer with their ‘intrinsic’ classification: luminal cancers, which are almost all ER positive, express cytokeratin 8 and 18 typically for the mammary gland, and are divided into luminal A, which are mostly histologically low grade, and luminal B, which tend to be of high grade with a worse prognosis; HER2-positive cancers, which show amplification and overexpression of the ERBB2 gene, do not express hormone receptors and are of poor prognosis; basal-like breast cancers, which overlay markedly with ER-, PgR-, and HER2-negative (triple negative) tumors with a poor prognosis and expression of cytokertatins of the basal epithelial layer (CK 5/6, CK 17; Sørli et al. 2001, 2003, Sotiriou et al. 2003). As indicated, these subgroups correspond closely to the earlier classification on the basis of hormone receptor and HER2 status.

Several groups have grasped the challenge to develop genomic tests based on genomic profiling with the expectation that this might better predict for clinical outcome than the standard pathological and clinical markers (Table 2). A 70-gene signature called MammaPrint (Agendia, Amsterdam, The Netherlands) has been developed using frozen samples from a group of 78 patients. These were selected retrospectively with node-negative breast cancer smaller than 5 cm, no adjuvant chemotherapy, and younger than 55 years (van ’t Veer et al. 2002). The top 70 genes, which most significantly correlated with clinical outcome (distant metastases within 5 years), were shown to accurately classify tumors in good or poor prognostic categories. The signature was validated retrospectively on a set of 295 patients including both lymph node positive and negative disease. This validation showed that the gene
signature outperformed all of the traditional clinical prognostic factors (van de Vijver et al. 2002). Unfortunately, some of the patients had received adjuvant systemic therapy, and some were also included in the training set. A second validation study using the TRANSBIG series also confirmed the prognostic capacity of the gene signature (Buyse et al. 2006) and a comparison with the Adjuvant! Online software revealed that the genomic test appeared to be able to more accurately predict outcome in discordant cases (Sotiriou & Pusztai 2009). The FDA has approved MammaPrint for clinical use as a prognostic test for women of the age 61 years or below with either ER-positive or ER-negative, lymph node-negative breast cancer, but a prospective validation trial to determine its clinical utility is currently recruiting. This is the microarray in node-negative disease may avoid chemotherapy (MINDACT) trial, which is expected to enrol 6000 breast cancer patients who will have a risk assessment using clinicopathological factors (Adjuvant! Online) and the 70-gene profile. Patients will receive adjuvant chemotherapy if both the tests predict high risk and if both indicate a low risk chemotherapy will be withheld. In case of discordant results, patients will be randomized to follow one result (Bogaerts et al. 2006).

The Oncotype DX signature was designed to predict better the risk of distant recurrence in patients with ER-positive early breast cancer receiving tamoxifen. This test is based on real-time PCR measurement of the expression of 16 genes with known significance in breast cancer and of 5 reference genes (Table 3). A recurrence score (RS) is calculated with a mathematical algorithm, which was developed and established using samples of the tamoxifen arms of the NSABPB20 and B-14 trials (Paik et al. 2004). The RS is a continuous measurement of risk, but it has generally been used to identify three groups with a low, intermediate, and high risk of distant recurrence, which were associated with <10%, 10–30%, and >30% 10-year recurrence rates among tamoxifen-treated patients. The RS has been described to be predictive of overall survival and distant recurrence independent of age and tumor size.

**Table 2 Multigene parameters in breast cancer**

<table>
<thead>
<tr>
<th>Gene signature</th>
<th>Number of genes assessed</th>
<th>Tissue</th>
<th>Application</th>
<th>Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>MammaPrint</td>
<td>70</td>
<td>Fresh frozen</td>
<td>Prognostic for recurrence within 5 years in all node-negative and node-positive patients</td>
<td>MINDACT</td>
</tr>
<tr>
<td>Oncotype DX</td>
<td>21</td>
<td>FFPE</td>
<td>Residual risk of DR in ER-positive patients treated with tamoxifen or AIs; and predictive of chemotherapy benefit in node-negative ER-positive patients</td>
<td>TAILORx</td>
</tr>
<tr>
<td>Genomic-grade index</td>
<td>97</td>
<td>Originally fresh frozen, validated for FFPE</td>
<td>Prognostic, prediction of relapse in endocrine-treated ER-positive breast cancer</td>
<td></td>
</tr>
<tr>
<td>Molecular grade index</td>
<td>5</td>
<td>FFPE</td>
<td>Predicts poor outcome despite endocrine therapy in ER-positive breast cancer</td>
<td></td>
</tr>
<tr>
<td>Rotterdam signature</td>
<td>76</td>
<td>Fresh frozen</td>
<td>Prognostic for development of distant metastases within 5 years</td>
<td></td>
</tr>
</tbody>
</table>

FFPE, formalin-fixed, paraffin-embedded; ER, estrogen receptor; DR, distant recurrence; AI, aromatase inhibitor.

**Table 3 A 21-gene panel used for the Oncotype DX assay to calculate the recurrence score: 16 cancer-related genes and 5 reference genes**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proliferation</td>
<td>Ki67</td>
</tr>
<tr>
<td></td>
<td>STK15</td>
</tr>
<tr>
<td></td>
<td>Survivin</td>
</tr>
<tr>
<td></td>
<td>Cyclin B1</td>
</tr>
<tr>
<td></td>
<td>MYBL2</td>
</tr>
<tr>
<td>Estrogen</td>
<td>ER</td>
</tr>
<tr>
<td></td>
<td>PgR</td>
</tr>
<tr>
<td></td>
<td>BCL2</td>
</tr>
<tr>
<td></td>
<td>SCUBE2</td>
</tr>
<tr>
<td></td>
<td>Stromelysin 3</td>
</tr>
<tr>
<td></td>
<td>Cathepsin L2</td>
</tr>
<tr>
<td>Invasion</td>
<td>GRB7</td>
</tr>
<tr>
<td>HER2</td>
<td>HER2</td>
</tr>
<tr>
<td>Other</td>
<td>GSTM1</td>
</tr>
<tr>
<td></td>
<td>CD68</td>
</tr>
<tr>
<td></td>
<td>BAG1</td>
</tr>
<tr>
<td>Reference</td>
<td>β-Actin</td>
</tr>
<tr>
<td></td>
<td>GADPH</td>
</tr>
<tr>
<td></td>
<td>RPLPO</td>
</tr>
<tr>
<td></td>
<td>GUS</td>
</tr>
<tr>
<td></td>
<td>TFRC</td>
</tr>
</tbody>
</table>
In the largest series to date, our group confirmed the performance of the RS in patients receiving an AI or tamoxifen. Importantly, these data revealed that the prognostic information in the RS and that of clinical features (e.g., nodal status, tumor size and grade) as integrated by Adjuvant! Online were almost entirely independent of one another (Dowsett et al. 2010; Fig. 3). This provides the opportunity to derive a single algorithm which integrates the two sets of features and will be more accurate than either alone. A high RS has been found to predict for chemotherapy benefit in hormone receptor-positive early breast cancer (Paik et al. 2006, Sparano & Paik 2008). To evaluate this further, the trial assigning individualized options for treatment (TAILORx) is aiming to recruit more than 10,000 patients with lymph node-negative hormone receptor-positive breast cancer. Patients with an intermediate RS will be randomly assigned to receive adjuvant chemotherapy or not as well as to receive subsequent endocrine therapy (Sparano 2006).

A genomic-grade signature has been developed to define molecular features of tumor differentiation that might relate to progression and metastasis better than histological grade (Rhodes et al. 2004, Sotiriou et al. 2006). It consists of a 97-gene signature, which is able to discriminate grade 2 tumors into low and high genomic-grade subgroups with outcomes comparable to histological low- and high-grade tumors. The genomic-grade signature was evaluated on different datasets and was found to be better associated with outcome than established clinical parameters (Loi et al. 2007) and prediction of relapse under endocrine treatment (Desmedt et al. 2009). This signature accentuates the importance of differentiation and particularly proliferation genes in ER-positive breast cancer.

Many other gene signatures have been developed and have undergone validation. One of them is the breast cancer gene expression ratio test, which only measures the ratio of HOXB13 to IL17BR (Ma et al. 2006, 2008, Wang et al. 2007). A high mRNA expression ratio was associated with a high risk of recurrence in tamoxifen-treated patients. Recently, the accuracy of this test could be improved by including proliferation-associated genes of the molecular grade index (Ma et al. 2008), which is an RT-PCR assay consisting of five genes that are able to identify a subgroup of ER-positive patients with a worse outcome despite endocrine therapy. The Rotterdam 76-gene signature was created on the basis of predicting the development of metastatic disease within 5 years using an unselected patient cohort regarding age, tumor size, grade, and hormone receptor status (Wang et al. 2005, Loi et al. 2007).

These signatures are composed of different gene sets with few overlaying genes, but there is predominance of these associated genes with proliferation. The choice for the oncologist should be made on the basis of the clinical context (e.g., pure prognosis or prognosis in the presence of endocrine treatment; all ages or just young age) and on the biopsy material available (e.g., MammaPrint needs fresh tissue, but some others including RS can use fixed tissue).

**Circulating tumor cells and tumor-specific DNA**

There is currently a major effort to identify biomarkers which can be obtained with minimally invasive methods and persist beyond surgery. The existence of circulating tumor cells (CTCs) in the blood of cancer patients was first reported in 1869, but only in the last decade has molecular methodology made it possible to detect them reproducibly (Smith et al. 1991). In parallel, advances in immunohistochemistry made it possible to identify disseminated tumor cells in the bone marrow.

For breast cancer, a high CTC count at diagnosis of metastasis is described as being a significant negative prognostic factor, and if the number of CTCs does not decrease, patients are likely to progress under...
chemotherapy (Cristofanilli et al. 2004). The future use of CTC measurements is very probably to predict therapy efficacy and resistance after initial exposure to therapy, and may be beneficial in monitoring response to treatment. Although many reports on the significance of CTCs have been published, in 2007, the American Society of Clinical Oncology tumor marker group concluded that treatment decisions should not be influenced by CTC counts (Harris et al. 2007). A particularly attractive concept that is beginning to meet the expectations is the identification of specific biomarkers on the CTCs, e.g. HER2 (Fehm et al. 2005).

Another possible approach to improve simplifying breast cancer management is the study of circulating cell-free DNA (cfDNA). This may be either from nuclear or from mitochondrial origin. Increased levels have been detectable in several cancer types, and an association between nuclear cfDNA levels and malignancy as well as tumor size has been described (Catarino et al. 2008, Kohler et al. 2009). Others report the possibility to screen for PIK3CA mutations in cfDNA (Board et al. 2010). Such studies herald a day when the genetic aberrations of an individual’s tumor may be used to create absolute specificity for the cfDNA in their blood.

In addition, the occurrence of tumor-specific microRNA species in the tumor and blood of breast cancer patients has been investigated. Specific expression patterns have been described (Iorio et al. 2005, Mattie et al. 2006), and some were linked to clinicopathological variables (Heneghan et al. 2010), but at present further studies are needed to validate the specific function of microRNAs and their possible use as biomarkers.

Conclusions

The fact that breast cancer is not a uniform cancer entity but consists of several different subtypes with different molecular profiles, biological behavior, and risk profiles poses a challenge for the clinical management. Prognostic and predictive factors constitute important tools for the individualization of breast cancer therapy to provide efficient treatment and to spare patients with excellent low-risk profiles from unwanted side effects of overtreatment.

The established clinicopathologic markers, in particular ER and HER2, have clearly defined clinical applicability, but deficiencies in the methodologies of assessment may still affect their use. Additional tools are required to facilitate clinical decision-making processes especially for the optimal treatment of early hormone receptor-positive breast cancer.

Very few of the many individual prognostic markers evaluated are sufficiently powerful on their own to merit clinical use. Some of the multigene signatures are more powerful. Although some already gained the FDA approval, large prospective randomized trials are currently being conducted to evaluate their clinical applicability. A recent finding from our group that an IHC panel of just four frequently used markers (Cuzick et al. 2009) was at least as prognostic as the Oncotype RS indicates the need for further studies comparing new methodology with established, less-expensive methodology. We should not be misled by the seductiveness of the new.

Special attention needs to be paid to the design and conduct of clinical trials, which have the potential to validate emerging biomarkers for their clinical application. Careful assay design and validation is an absolute requirement alongside the collection of quality tissue and blood specimen to address the clinical question for which the marker has been developed and selected. If well designed and conducted, these trials have the potential to provide unique collections of clinical specimens, which may be of great use for future biomarker discovery and retrospective evaluation.

Despite the discussion of several emerging biomarkers in this paper, there are others that could have been included such as CYP2D6 polymorphism. However, there is controversy about the reliability of its association with response to tamoxifen because of a widely varying set of clinical reports.

Given the development of new targeted molecular therapies, there will continue to be a need for identifying and devising new markers that will be able to predict for specific response. It will be a challenge for scientists and clinicians to select the most promising ones particularly where overexpression of the target is not required for activity.

With the effort being exploited in this area and the enormous strides being made in characterizing the molecular characteristics of individual cancers, the future should provide us with unique case-specific patterns of biomarkers, which will help to optimize tailored therapies and individualize breast cancer patient care.

Declaration of interest

M T Weigel has no conflict of interest that could be perceived as prejudicing the impartiality of the research reported; M Dowsett receives research funding, honoraria for advisory boards, and lecture fees from AstraZeneca, Novartis, and Roche.
M T Weigel and M Dowsett: Biomarkers in breast cancer

Funding

M Dowsett is grateful for funding from Breakthrough Breast Cancer, Cancer Research UK and the NIHR Royal Marsden Hospital Biomedical Research Centre. M T Weigel is supported by Breakthrough Breast Cancer and the Deutsche Krebshilfe (grant number 108999).

References


Keyomarsi K & Pardee AB 1993 Redundant cyclin overexpression and gene amplification in breast cancer cells. PNAS 90 1112-1116. (doi:10.1073/pnas.90.3.1112)


Matthews J & Gustafsson JA 2003 Estrogen signaling: a subtle balance between ER alpha and ER beta. Molecular Interventions 3 281-292. (doi:10.1124/mi.3.5.281)


Skliris GP, Leygue E, Curtis-Snell L, Watson PH & Murphy LC 2006 Expression of oestrogen receptor-beta in...
oestrogen receptor-alpha negative human breast tumours. British Journal of Cancer **95** 616–626. (doi:10.1038/sj.bjc.6603295)


Low molecular weight cyclin E is specific in breast cancer and is associated with mechanisms of tumor progression. *Cell Cycle* **8** 1062–1068.
