Breast cancer prognostication and prediction: are we making progress?

P. E. Lønning*

Section of Oncology, Institute of Medicine, University of Bergen, and Department of Oncology, Haukeland University Hospital, Bergen, Norway

Currently, much effort is being invested in the identification of new, accurate prognostic and predictive factors in breast cancer. Prognostic factors assess the patient's risk of relapse based on indicators such as intrinsic tumor biology and disease stage at diagnosis, and are traditionally used to identify patients who can be spared unnecessary adjuvant therapy based only on the risk of relapse. Lymph node status and tumor size are accepted as well-defined prognostic factors in breast cancer. Predictive factors, in contrast, determine the responsiveness of a particular tumor to a specific treatment. Despite recent advances in the understanding of breast cancer biology and changing practices in disease management, with the exception of hormone receptor status, which predicts responsiveness to endocrine treatment, no predictive factor for response to systemic therapy in breast cancer is widely accepted. While gene expression studies have provided important new information with regard to tumor biology and prognostication, attempts to identify predictive factors have not been successful so far. This article will focus on recent advances in prognostication and prediction, with emphasis on findings from gene expression profiling studies.

Key words: breast cancer, microarray, nodal status, predictive, prognostic

introduction

Breast cancer is the most commonly occurring malignancy in women and is responsible for approximately 500 000 deaths per year worldwide [1]. In 2002, 1.15 million new cases of breast cancer were diagnosed [2]. Despite the rising incidence of this disease, survival rates have improved in recent years in many countries due to the encouraging trend towards earlier detection, together with increasing use of systemic adjuvant treatments.

Prognostic factors are important for forecasting outcomes in individual patients and can be used to help refine treatment choices. Nodal status (positive or negative) is the most important parameter used to define risk category in early breast cancer. Although imperfect, the strength of nodal status as a prognostic factor relates to the fact that it is a 'pure' prognostic factor and does not also predict for response to therapy [3]; that is, nodal status does not affect response to systemic therapy [4]. To date, the presence of estrogen receptors (ERs) and/or progesterone receptors is the only well-defined predictive factor, and predicts for responsiveness to endocrine therapy. With respect to chemotherapy, no single predictive factor has yet been universally accepted.

In theory, the identification, validation and application of suitable predictive and prognostic factors will help to ensure that only those patients likely to benefit will receive a given treatment. Moreover, identification of genes involved in responsiveness to therapy may lead to the characterization of new therapeutic targets and the subsequent availability of more treatment options for patients with resistant disease. The optimal application of prognostic and predictive factors is depicted in Figure 1. This article will consider the ability of emerging predictive and prognostic factors to reduce the overall treatment burden.

prognostic factors

Traditionally, the dogma has been that prognostic factors help physicians to determine which patients with breast cancer need adjuvant therapy, whereas predictive factors indicate which adjuvant therapy is most appropriate. The most important benefit of prognostic classification may be to help physicians identify patients in whom adjuvant therapy could be avoided, thus sparing these patients from treatment-related side effects.

Nodal status, the major factor defining risk category, may be considered a 'pure' prognostic factor, as patients with multiple lymph-node metastases can achieve similar benefits from therapy, in terms of relative reduction in the risk of relapse, as those with few or no lymph-node metastases [3, 4]. Nodal status is well accepted as a useful prognostic factor, with node-negative status (including the sentinel-node-negative classification) identifying patients with low-risk disease [5]. However, the St Gallen treatment guidelines were updated in 2005 to include the new, intermediate-risk category, so that the presence of positive axillary nodes in the absence of additional
high-risk characteristics no longer defines high-risk disease [6].

Two new adverse prognostic factors were also accepted: HER2/neu overexpression or amplification, and peritumoral vascular invasion. Further details of the St Gallen risk classification are provided in Table 1 [6]. Importantly, at the 2005 St Gallen meeting, for the first time, endocrine responsiveness, not risk category, was identified as the primary consideration when making treatment decisions [6].

**identification of novel prognostic factors**

Prognostic factors in breast cancer have been widely investigated, and more than 100 individual factors have been reported in the literature. However, few of these factors have found their way into clinical application as prognostic tools, or contributed greatly to our understanding of tumor biology.

Although several novel prognostic factors, including lymph node micrometastases and PAI-1 expression [7, 8], have been proposed, their clinical value remains uncertain at this stage. It should be noted that no new prognostic factor should be considered independent of its potential predictive role. On the one hand, there is little benefit in selecting patients for a given therapy based on the presence of a factor defining poor prognosis, if the same factor also defines low sensitivity to the therapeutic agent to be given. On the other hand, and of even greater concern, is that patients who are considered to have low-risk disease, but who have tumors that are highly responsive to a given therapy, may not be offered the treatment on the basis of prognosis alone, even though it could substantially improve their long-term outcome [4]. This may be illustrated using two conventional prognostic factors, TP53 mutational status and HER2/neu amplification. While both parameters identify patients with a poor prognosis [9–12], both are also associated with (although not fully predictive for) responsiveness to specific treatments [13–16].

**genomic signatures and multigene RT–PCR as prognostic tools**

Gene profiles with reported prognostic value have been identified using different microarray techniques, but the potential predictive component of these profiles remains unclear. In general, gene profiles have been identified through one of two statistical methods: hierarchical clustering and supervised analysis [17].

Perou et al. used hierarchical clustering to characterize variation in gene expression patterns between different breast tumors [18]. Large variation was found in expression patterns

---

**Table 1.** Risk categories for patients with operable breast cancer [6]

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Disease/patient characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Node-negative plus all of the following</td>
</tr>
<tr>
<td></td>
<td>Pathological tumor size ≤2 cm</td>
</tr>
<tr>
<td></td>
<td>Tumor grade 1</td>
</tr>
<tr>
<td></td>
<td>No peritumoral vascular invasion</td>
</tr>
<tr>
<td></td>
<td>HER2/neu gene neither over-expressed nor amplified</td>
</tr>
<tr>
<td></td>
<td>Patient age ≥35 years</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Node-negative plus at least one of the following</td>
</tr>
<tr>
<td></td>
<td>Pathological tumor size ≥2 cm</td>
</tr>
<tr>
<td></td>
<td>Tumor grade 2–3</td>
</tr>
<tr>
<td></td>
<td>Peritumoral vascular invasion</td>
</tr>
<tr>
<td></td>
<td>Confirmed HER2/neu gene over-expression or amplification</td>
</tr>
<tr>
<td></td>
<td>Patient age &lt;35 years</td>
</tr>
<tr>
<td></td>
<td>OR</td>
</tr>
<tr>
<td></td>
<td>Node-positive (1–3 nodes) plus</td>
</tr>
<tr>
<td></td>
<td>HER2/neu gene neither over-expressed nor amplified</td>
</tr>
<tr>
<td>High</td>
<td>Node-positive (1–3 nodes) plus</td>
</tr>
<tr>
<td></td>
<td>HER2/neu gene over-expression or amplification</td>
</tr>
<tr>
<td></td>
<td>Node-positive (≥4 nodes)</td>
</tr>
</tbody>
</table>

*Histological and/or nuclear grade.
from individual tumors, reflecting the great diversity of breast cancer as a disease. Furthermore, tumors could be classified into different phenotypic subtypes based on differences in their gene expression profiles, and a hierarchical order was identified: luminal subtypes A, B and C; normal breast-like; basal-like; and HER2+ tumors [18–20]. Further analyses revealed the prognostic power of the subgroups identified by Perou et al. [18, 19]. However, it has recently been shown that neither hierarchical clustering nor supervised analysis allows predictive profiles to be generated from the same material [21].

In contrast, van’t Veer et al. used supervised analysis to identify a 70-gene ‘poor prognosis’ signature, which was strongly predictive of a short time to distant metastases. [22]. The prognostic power of this ‘Amsterdam’ 70-gene signature was confirmed in a study of 295 patients who were classified as having either poor or good prognosis. The results showed the Amsterdam profile to be a more powerful predictor of disease outcome than standard systems based on clinical and histological criteria [23]. Importantly, the prognostic signature was found to predict outcome among patients with node-negative disease (who are not generally given adjuvant treatment) and those with node-positive disease (the majority of whom will receive chemotherapy). However, the number of patients included and the design of the study may not fully exclude a predictive component of this signature. The prognostic value of the Amsterdam 70-gene signature has since been independently confirmed [24]. Alternative prognostic signatures have recently been identified, including the ‘Rotterdam’ 76-gene profile [25]. However, the existence of a unique prognostic profile has been brought into question by others [26], after the demonstration that a ‘family’ of different profiles could be generated from the data set used by the Amsterdam group.

Histological grading has been used for many years to assign risk categories, but only half of all breast cancers are classified as grade 1 (low-risk) or 3 (high-risk), the remainder being classified as grade 2. Recently, attention has turned to genomic grading to help refine histological grading, specifically grade 2, which, when used alone, has little prognostic value. Using microarray analysis, the Brussels Genomic Grade 97-gene signature was identified, which could accurately reclassify patients with histologic grade 2 tumors into those with a high or low risk of recurrence, thereby improving the prognostic value of tumor grading [27].

Several other prognostic gene signatures have been identified using microarray analysis. For example, a ‘stem cell’ 11-gene signature described by Glinsky and colleagues [28], and the 21-gene profile [32] and the 2-gene expression ratio [33] reported to forecast outcome in tamoxifen-treated patients. Gene expression profiling studies are becoming increasingly common, and the St Gallen expert consensus guidelines overwhelmingly endorse further prospective studies to identify potential predictive and prognostic factors [6]. Further studies are required to ascertain the impact of gene profiling on patient treatment.

**predictive factors**

**predictors of response to chemotherapy**

While the identification of reliable predictive factors has the potential to spare patients from ineffective treatment and unnecessary side effects, the reverse—that a factor may guarantee therapeutic success—may be more difficult to achieve. Thus, while ER negativity is associated with lack of response to endocrine manipulation, not all patients with ER+ tumors may benefit from such therapy. Similarly, while the absence of HER2/neu overexpression has been established as a predictive factor for non-responsiveness to trastuzumab therapy, not all HER2/neu-overexpressing tumors are trastuzumab sensitive [34], reflecting the complexity of breast cancer genetics.

A second issue relates to whether a predictive factor may be a causal factor or just a co-variate. Whereas early studies suggested that HER2 overexpression or amplification of the HER2/neu gene predict for sensitivity to anthracyclines [35, 36], recent studies have shown that TOPO-IIα, not HER2, overexpression predicts for anthracycline sensitivity in tumors with coamplification of the two genes [37–41]. The ongoing ‘Trial Of Principle’ will prospectively test the value of TOPO-IIα in predicting the efficacy of anthracycline therapy in women with hormone-receptor-negative breast cancer [42].

Mutations in TP53 and overexpression of p53 (the protein encoded by the TP53 gene) have also been shown to correlate with chemosensitivity in patients with breast cancer. Specifically, responsiveness to anthracycline- or mitomycin-containing chemotherapy is reduced by defective TP53 status. In contrast, the efficacy of paclitaxel seems independent of p53 expression [13–15, 43]. However, available data from trials do not support the use of tumor TP53 status when selecting patients for a given treatment. Firstly, the number of patients in
these trials was limited. Secondly, the predictive power of mutant TP53 is not by itself sufficient for therapy selection; tumors harboring mutant p53 may still be responsive to anthracycline therapy, while some tumors with wild-type p53 are resistant to therapy, indicating that additional factors are involved in the control of chemosensitivity. Furthermore, every gene product, including p53, works as part of cellular ‘cascades’, which include upstream ‘activators’ and downstream ‘executors’. Thus, while epigenetic silencing and somatic mutations in the p21 gene (which acts downstream of p53) have been excluded as a cause of drug resistance [44], the TP53 upstream activator, chk2, is subject to multiple splicing, and some of these splice variants may influence gene function [45]. Thus, more research is necessary before TP53 status can be accepted as a predictive tool for breast cancer therapy.

gene expression profiles as predictive factors
Several ongoing trials are using microarray analysis and gene expression profiling with the aim of identifying better predictive factors for response to chemotherapy. As microarray techniques allow for the simultaneous analysis of multiple genes, in theory, this should be a promising approach for identifying multiple factors acting in concert to influence response to therapy. Gene profiles revealing variable association to chemosensitivity have been described by several groups [21, 46–51]. Interestingly, the National Surgical Adjuvant Breast and Bowel Cancer Project (NSABP) group has further explored their 21-gene signature, previously shown to have prognostic value in women with hormone-receptor-positive disease taking adjuvant tamoxifen [32], with respect to cyclophosphamide, methotrexate and 5-fluorouracil (CMF) treatment. In this study, it was shown that the 21-gene signature also has predictive value for responsiveness to CMF in the same group of patients [52].

Critical examination of these studies reveals that each profile was developed and validated in a limited number of patients. For most studies, the sensitivity and specificity reported to date does not support their use in therapy selection, and none of these profiles have been accepted for general use in the clinic. While ongoing studies are assessing potential predictive profiles for chemotherapy as well as endocrine therapy, others have questioned the theoretical basis for developing such predictive profiles [53–55]. In a recent study, gene expression profiles obtained from tumors in patients undergoing neoadjuvant chemotherapy failed to correctly classify patients with chemoresistant disease at high risk of disease progression [21], but the same classification clearly outlined prognosis in the same cohort [19].

conclusions
Current breast cancer treatment guidelines are based on clinical trial evidence obtained in defined patient populations, and treatment algorithms are developed by relating clinical trial findings to specific patient subgroups. There is a general consensus that better prognostic and, in particular, predictive factors are needed to assist in treatment decision-making on an individual patient basis.

Considering prognostication, gene profiling seems promising, although further validation, particularly with respect to potential ‘predictive interactions’, is mandatory. For predictive testing, emerging evidence suggests TOPO-II amplification or deletions may be appropriate factors for selecting patients for anthracycline dosing. Gene expression profiling may become important as a tool to define predictive factors; however, to achieve this goal, statistical approaches analyzing gene expression profiles, based on functional hypotheses about gene networks, will be required [4, 56, 57].

disclosures
Professor Lønning has reported no financial relationships with companies whose products are mentioned in this supplement.

references


