1. Introduction

Breast cancer is the most common malignancy and the leading cause of cancer death in women globally. There has been a sharp increase in its incidence especially in the developed world due to a combination of better detection and lifestyle changes. Breast cancer is a disorder influenced by genetic, environmental, behavioral, and reproductive factors. The most significant risk factors are gender and age. Hereditary forms of breast cancer are often related to mutations in two high-penetrance susceptibility genes namely BRCA-1 and BRCA-2 (1), and account for around 5% of all breast cancer cases. Women who are born with these mutations have 10–30-fold increased risk of developing breast cancer compared to the general population and a cumulative lifetime risk of 60–80%. Sporadic forms of breast cancer account for around 95% of cases and are a consequence of somatic mutations acquired over the lifetime; they appear to be in part related to polymorphisms in low-penetrance genes that encode proteins involved in DNA repair, cell signaling pathways, estrogen metabolism, etc. (2, 3). In the last few decades, the survival rate of breast cancer has improved due to advances in mammography and adjuvant therapy.

Histopathologically identical tumours may exhibit different biological behaviors in terms of severity, course, and response to therapy, reflecting disease heterogeneity; in addition, variability of the host immune response further contributes to differences in treatment outcomes, underscoring the need for better understanding of this disease and its relation to the host (4). At the biological level, breast cancer is a complex disease caused by multiple genetic and epigenetic alterations that ultimately lead to changes in cell processes, including cell proliferation, apoptosis, and angiogenesis, with subsequent acquisition of a malignant phenotype (5). The main genetic abnormalities that are observed include increased proto-oncogene expression, inactivation of tumour suppressor genes, chromosomal instability,
alterations in DNA repair genes, telomerase reactivation, and epigenetic changes, resulting in dysregulation of cell proliferation, clonal selection, and tumour formation (6). As such one can expect breast cancer to be a heterogeneous disease and better prognostic and predictive biomarkers are clearly needed to better manage this disease.

The treatment of breast cancer continues to be challenging because of the heterogeneity of the disease. Breast cancer is staged by the TNM classification that assigns tumours to different stages based on depth of tumour invasion and presence of nodal and distant metastases. However, considering the heterogeneity in outcome of patients diagnosed with equivalent TNM stage, this classification system is suboptimal in tumour characterization or prognostication. In early-stage breast cancer, several clinicopathological factors are used to refine prognostication over and above TNM staging. These factors include histological grade, lymphovascular invasion and estrogen receptor (ER)/progesterone (PR) status. Some of these factors have been incorporated into algorithms such as Adjuvant! Online to estimate the individual risk of cancer relapse (7-9). More recently, amplification and/or overexpression of the human epidermal growth factor receptor 2 (Her2), a therapeutic target, has been associated with worse prognosis, although its clinical utility as a prognostic marker remains uncertain (10-12). The variation in clinical outcome despite similar clinical and pathological prognostic scores seriously compromises the ability to advise women in making fully informed decisions about adjuvant systemic treatment after definitive surgery.

Over the past few decades, substantial effort has been invested in the identification and validation of prognostic markers over and above ER, PR and Her2, in an attempt to improve risk stratification for breast cancer. As the evaluation of candidate prognostic markers is often limited by inadequate study design and analyses, formal recommendations for reporting tumour marker prognostic studies have been suggested, including guidelines on assay methods, study design and data analysis (13).

In recent years, gene expression microarray-based technology has resulted in the identification of breast cancer molecular subtypes and gene-expression prognostic signatures (14-16). These classification and prognostic expression signatures hold great promise, but there are concerns regarding their significance independent of ER/PR status (17, 18). The process of validating the clinical utility of two such prognostic gene expression signatures, Oncotype DX and Mammaprint, is ongoing through the TAILORx and MINDACT trials respectively. Until these are validated prospectively, the increasing usage of these two profiling tests is unfortunately based on mainly retrospective data.

A fair proportion of breast cancers cannot be adequately resected upfront. In these situations, neoadjuvant chemotherapy is often given first. In recent times, even those tumours that are borderline resectable are often treated with neoadjuvant chemotherapy in an attempt to improve cosmetic results. Neoadjuvant chemotherapy provides prognostic information as the achievement of pathologic complete response (pCR) is associated with prolonged survival. The increasing use of neoadjuvant chemotherapy in patients with primary breast cancer makes it important to develop predictive markers of pCR, which is a surrogate marker of improved survival (19, 20). In addition, neoadjuvant chemotherapy allows the biological effect of the therapy to be evaluated as the surgically resected tumour
after treatment can be examined and compared with the pre-treatment biopsy sample, providing an opportunity to study tumour biology \textit{in vivo}. Neoadjuvant trials thus offer an excellent opportunity to study tumour DNA, RNA and protein changes and to evaluate new prognostic and predictive biomarkers of treatment response. In addition, it has the potential to reveal post-treatment biomarkers that could be complementary or even superior to the routine baseline biomarkers currently in use.

This chapter will review protein biomarkers in breast cancer. It will focus on established biomarkers, the timepoints of obtaining biomarkers, and the type of specimens on which to analyze these biomarkers. The different methods of measuring such biomarkers will also be described. In addition, several candidate protein biomarkers (e.g., Topo2α, serum Her2, Cox2, MGMT, Hsp-70) will be reviewed for their possible utility as post-chemotherapy markers.

### Body

<table>
<thead>
<tr>
<th>Established and/or clinically relevant biomarkers</th>
<th>Potentially clinically relevant biomarkers</th>
<th>Biomarkers under evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER(^*)</td>
<td>Topo2α(^*)</td>
<td>EGFR</td>
</tr>
<tr>
<td>PR(^*)</td>
<td></td>
<td>p53</td>
</tr>
<tr>
<td>Her2(^*)</td>
<td></td>
<td>Bcl2</td>
</tr>
<tr>
<td>Ca15-3(^*)</td>
<td></td>
<td>MGMT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hsp-70(^*)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>COX2(^*)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Osteopontin</td>
</tr>
</tbody>
</table>

*Table 1. Biomarkers in breast cancer (*includes serum. \(^*\)predictive of response.*)

**A. Biomarkers in breast cancer**

1. **Prognostic versus predictive biomarkers**

A prognostic biomarker provides information about the patient’s overall cancer outcome, regardless of therapy, whilst a predictive biomarker gives information about the effect of a therapeutic intervention. A predictive biomarker can also be a target for therapy. Among the genes and proteins that have proven to be of relevance in cancer are well known predictive markers such as ER, PR and Her2/neu in breast cancer, c-KIT mutations in gastrointestinal stromal tumours, EGFR mutations in non-small cell lung cancer, and BCR-ABL fusion protein in chronic myeloid leukaemia.

2. **Criteria of candidate biomarkers**

Several factors are important in selection and validation of candidate biomarkers. The analysis platform must be sufficiently robust to detect subtle changes between tumours.
Sample sets must be robust enough to reduce pre-analytical data biases and must reflect the intended use of the marker or marker set. Independent sample sets must be used to validate the prognostic and predictive power of biomarkers particularly when many biomarkers are assessed on small sample sets. Lastly, bioinformatics support is essential at all steps in any project. In addition, these markers would need to be validated, usually retrospectively first in existing large clinical datasets and ultimately in prospective randomized trials.

3. Gene expression biomarkers in breast cancer

The complementary DNA (cDNA) -microarray technology has made it possible to analyze the mRNA expression of numerous genes simultaneously to better characterize breast cancers, including classification and prognostication. Several studies using transcriptional profiling have classified breast cancer into different subtypes with implications in patient prognosis (21-23), frequency of genomic alterations (24, 25), and therapy response (26, 27). In breast cancer classification, the first tier of separation is between ER-negative and ER-positive tumours. Five breast cancer molecular subtypes have been identified using this technology, of which the luminal (A and B) type is ER-positive and accounts for 60% of breast tumours; the Her2 overexpressing type accounts for 15–20%; the ER and Her2 negative basal-like type accounts for 20% of the cases and has a guarded prognosis; and lastly the normal-like type, which has no definitive clinical value (28, 29). ER-positive tumours respond to endocrine therapy (e.g., tamoxifen, aromatase inhibitors), and Her2 positive tumours are eligible for targeted therapy with trastuzumab or lapatinib, whereas the basal-like type has a more aggressive phenotype and while generally responsive to various chemotherapy regimens tends to acquire resistance quickly and has short survival (28-31). Currently, the advocates of this classification have suggested that the normal-like subtype might actually be an artefact of sample representation, that is, contamination of the mammary tissue by normal cells (32, 33). More recently, three other ER-negative subtypes have been described, the molecular apocrine tumour, the interferon, and lastly the claudin-low, which expresses breast epithelial stem cell markers. However, a definition of their clinical significance is still needed (34). Despite its significant contribution, the ‘gene signature’ described above is not a definitive classification method, but rather a developing work model that needs to be refined, considering that more subtypes have been described (5). Prognostic gene expression signatures in the form of Oncotype DX and Mammaprint have been tested in various large clinical datasets retrospectively to show prognostic value as well as value in predicting benefit from adjuvant chemotherapy, and are already in clinical use. On the other hand, predictive gene expression signatures for response to specific drug or drug regimens are still largely investigational although there have been many studies. This is because of small sample size in most studies, lack of independent validation sets in some studies, heterogeneity of the study population, a great variety of chemotherapy regimens that were evaluated in different studies, and variation in definition of response endpoint, making it difficult to pool the study results (35).

4. Protein biomarkers in breast cancer

Cancer arises from successive genetic changes, by which several cellular processes, including growth control, senescence, apoptosis, angiogenesis, and metastasis, are altered...
Serial Changes in Expression of Proteins in Response to Neoadjuvant Chemotherapy in Breast Cancer

Consequently, researchers initially searched for markers by employing genomic and transcriptomic approaches, providing new biomarkers (14, 37) and expanding our insight into the genetic basis of cancer. It is however currently understood that genetic analysis alone is insufficient. Alternative splicing of mRNA combined with numerous unique post-translational protein modifications can give rise to multiple protein species (38). Hence, compared to the genome, the proteome can provide a more dynamic and accurate reflection of both the intrinsic genetic programme of the cell and the impact of its immediate environment (39).

Since proteins are the effectors of cellular behavior, interrogation of the functional proteome is likely to complement data derived from transcriptional profiling. Thus, the integrated study of the expression and activation of multiple proteins and signaling pathways has the potential to provide powerful classifiers and predictors in breast cancer (40, 41). Currently, gene-profiling technology generally requires fresh or frozen tumor tissue (other than Oncotype DX), and is cumbersome and logistically demanding, which may limit its suitability for routine use in clinical practice for some time. As such, reliable protein markers that may be readily tested on routinely available biological specimens may be more widely applicable in the clinic.

5. Established protein biomarkers in breast cancer
   i. Estrogen Receptor (ER)/ Progesterone Receptor (PR)

Assays for tumour expression of ER and PR have established utility in the clinical management of patients with both early stage and advanced breast cancer. They are routinely obtained on all tumour specimens and immunohistochemistry (IHC) is the predominant method for measuring ER and PR in clinical practice. Receptor positivity (staining of cell nuclei is considered positive) is an important indicator of hormone responsiveness and identifies tumours for which endocrine therapy is a valuable therapeutic option in both the adjuvant and advanced disease setting. Expression of ER and/or PR within tumours correlates well with low histologic grade especially in postmenopausal women. Reports have highlighted the extent of variability in ER and PR IHC assay caused by a variety of factors including differences in specimen handling, tissue fixation, antigen retrieval, and antibody type. In addition, variability in interpretation of assay results is caused by different laboratory threshold values for positive and negative. These variations have resulted in serious issues with ER reliability. In view of the controversy over what constitutes a positive test, most laboratories will report the actual percentage of positive cells. While many agree that ≥5% is considered positive, tumours with a lower percentage (1-4%), or even no staining, may show a borderline response to endocrine therapy. The American Society of Clinical Oncology (ASCO) Tumour Marker Panel in 1995 concluded that: (1) ER and PR should be measured on every primary breast cancer and metastatic lesion if it would influence treatment planning, (2) ER and PR positivity supports use of endocrine therapy regardless of menopausal status in both adjuvant and metastatic disease, and (3) ER and PR receptors are weak prognostic indicators and should not be used to determine whether to treat a patient with adjuvant therapy. Newer guidelines from a
joint panel of the ASCO and the College of American Pathologists (CAP) provide recommendations to improve test accuracy and reporting of results (42). Of note, the panel now recommends that ER and PR assays be considered positive if there are at least 1% positive tumour nuclei in the sample on testing in the presence of expected reactivity of internal (normal epithelial elements) and external controls.

ii. Human epidermal growth factor receptor 2 (Her2)

Her2 is a proto-oncogene that encodes the production of Her2, a cell surface protein important in cell regulation. Abnormalities of Her2 occur in 25-30% of breast carcinomas, especially those that are poorly differentiated, lymph node positive, hormone receptor negative, flow aneuploid and/or show high proliferation rates. Her2 amplification and protein overexpression can be detected with Fluorescent In-situ hybridization (FISH) and IHC, respectively, both of which can be performed on paraffin-embedded tissue. Maximum sensitivity can be achieved by using both methods. The presence of Her2 overexpression predicts for response to anti-Her2 therapy such as trastuzumab and lapatinib. In addition, many studies have shown a positive response effect with anthracyclines in Her2 positive breast cancer, although there have been some studies recently to dispute this (43). Assay for this molecular marker is warranted as a routine part of the diagnostic work-up on all breast cancers, since Her2 overexpression is of major value in selection of anti-Her2 therapy in these patients. The bulk of available evidence supports the view that Her2 overexpression is associated with a poor prognosis. However, the value of this information in clinical practice is questionable, and guidelines from an expert panel on tumour markers in breast cancer convened by ASCO recommended against the use of Her2 in assessing prognosis (44). Given the substantial benefit of adjuvant trastuzumab in patients with Her2-overexpressing tumours, it is difficult to separate out the prognostic versus predictive utility of Her2.

Whilst the detection of tumour Her2 overexpression or amplification by IHC or FISH is standard clinical practice, the detection of serum (soluble) Her2 is a more controversial issue. In order to understand the relevance of serum Her2 we have to look at the structure of the Her2 protein. The Her2 protein is a 185-kDa transmembrane tyrosine kinase receptor with three defined domains: the intracellular tyrosine kinase portion, a short transmembrane portion, and the extracellular domain (ECD). The 105-kDa ECD (serum Her2) can be cleaved from the surface by metalloproteases and detected in the peripheral blood (45). It has been reported that trastuzumab inhibits Her2 extracellular domain cleavage; this is important considering that the remaining cleaved HER2 receptor is constitutively activated (46, 47), suggesting that the detection of sHer2 also reflects a biologic process leading to a more aggressive tumour behavior (48). Elevated levels of sHer2 have been observed in patients with primary (49) or metastatic breast cancer (50, 51). As detailed below in the specific biomarker section (E.3), there are some studies to suggest that elevated serum Her2 levels are a negative prognostic and predictive factor.
iii. Ca15-3

Ca15-3 detects circulating MUC-1 antigens in the blood. There are several studies that support the prognostic utility (52-55) of MUC-1 in early stage breast cancer. The trials showed as common finding that Ca15-3 was prognostic of disease free survival either on uni-variate or multi-variate analysis. We however do not use Ca15-3 to monitor patients with early stage breast cancer because there is no impact on the decision of chemotherapy regimen as established in a prospective clinical trial. In fact the sole approved use of this test (as per ASCO guidelines) is to monitor response to therapy in the metastatic breast cancer setting.

6. Candidate Protein Biomarkers with possible clinical application in breast cancer
i. Topo2-alpha (Topo2α)

DNA topoisomerase 2-alpha is an enzyme that in humans is encoded by the Topo2α gene. This gene encodes a DNA topoisomerase, an enzyme that controls and alters the topologic states of DNA during transcription. This nuclear enzyme is involved in processes such as chromosome condensation, chromatid separation, and the relief of torsional stress that occurs during DNA transcription and replication. It catalyzes the transient breaking and rejoining of two strands of duplex DNA which allows the strands to pass through one another, thus altering the topology of DNA. There is increasing interest on Topo2α and anthracycline sensitivity, although the results in the past have been somewhat mixed. The BCIRG006 investigators (56) have appropriately looked for markers of benefit from anthracyclines and have suggested in a large subset analysis that Topo2α co-amplification along with Her2 amplification could indicate a subset of patients who definitely benefit from anthracyclines, and, conversely, that the majority of patients who lack Topo2α co-amplification might possibly be just as well treated with trastuzumab without anthracyclines. However, because there is no widely available and validated Topo2α test and these data have not yet been corroborated independently, Topo2α testing is currently still not routinely performed in the clinic.

7. Protein biomarkers in breast cancer under evaluation
i. Epidermal Growth Factor Receptor (EGFR)

The epidermal growth factor receptor (EGFR; ErbB-1; Her1 in humans) is the cell-surface receptor for members of the epidermal growth factor family (EGF-family) of extracellular protein ligands. The epidermal growth factor receptor is a member of the ErbB family of receptors, a subfamily of four closely related receptor tyrosine kinases: EGFR (ErbB-1), Her2/neu (ErbB-2), Her3 (ErbB-3) and Her4 (ErbB-4). EGFR overexpression can be detected with IHC or FISH. In preclinical models of breast cancer, overexpression of EGFR leads to malignant transformation of mouse cells. It is associated with increased proliferation and resistance to apoptosis (57). One study analyzed 130 breast carcinomas using IHC analyses for the levels of nuclear and non-nuclear EGFR, and found that 37.7% of the cohort immunostained positively for nuclear EGFR and 6.9% had high levels of expression. More importantly, survival analysis
revealed a significant inverse correlation between high nuclear EGFR expression and overall survival. Furthermore, expression of nuclear EGFR correlated positively with increased levels of cyclin D1 and Ki-67, both of which are indicators for cell proliferation (58). The expression of EGFR and its association with shorter survival observed in this study has also been reported in other studies (59), although its routine use in breast cancer at this time is still controversial.

ii. p53

p53 (also known as protein 53 or tumour protein 53), is a tumour suppressor protein that in humans is encoded by the TP53 gene. p53 is crucial in multicellular organisms, where it regulates the cell cycle and, thus, functions as a tumour suppressor that is involved in preventing cancer. As such, p53 has been described as "the guardian of the genome" because of its role in conserving stability by preventing genome mutation. Mutations of the p53 gene cause variant p53 proteins to have an increased half-life. These variant p53 proteins accumulate in the cell and can be detected with IHC in about 90% of cases by increased nuclear staining. One study examined a chemoresistant subgroup of breast cancers (triple negative breast cancer) and showed that p53 was possibly prognostic (60). However, although over-accumulation of p53 protein has been associated with worse survival in breast cancer patients, it also correlates with cell proliferation and thus may not be an independent prognostic factor (61). In addition, the results of its prognostic significance in breast cancer have been inconsistent, and it is therefore not routinely used in breast cancer management.

iii. Bcl2

Expression of Bcl2, an anti-apoptotic protein, has been associated with low-grade, slowly proliferating, ER positive breast tumours (62, 63). In a report (64) which pooled five studies of 11,212 women with early-stage breast cancer together for analysis, individual patient data including tumour size, grade, lymph node status, use of adjuvant endocrine therapy and/or chemotherapy, and mortality were analyzed. Bcl2, ER, PR and Her2 levels were ascertained in all tumours. A Cox model was used to explore the prognostic significance of Bcl2. The study found that in univariate analysis, ER, PR and Bcl2 positivity was associated with improved survival and Her2 positivity with worse survival. Intriguingly, in multivariate analysis, Bcl2 positivity retained independent prognostic significance (hazard ratio 0.76). Bcl2 was a powerful prognostic marker in both ER negative (HR 0.63) and ER positive disease (HR 0.56), and in both Her2 negative (HR 0.55) and Her2 positive disease (HR 0.70), regardless of the type of adjuvant therapy received. The study also looked at the addition of Bcl2 to the Adjuvant! Online prognostic model, for a subset of cases with 10-year follow-up data and showed that Bcl2 improved the survival prediction.

8. Biomarkers elucidated by high throughput methods

Serum and plasma protein profiling studies by mass spectrometry (MALDI-TOF or SELDI-TOF) have yielded numerous protein peaks that are potentially diagnostic, prognostic, or
predictive in breast cancer. However, thus far, only a small percentage of reported peaks have been structurally identified. Moreover, since most studies did not investigate other cancer types or patients with benign breast disease, the specificity of reported markers for breast cancer still has to be addressed.

i. Diagnostic markers

The potential of proteomic pattern analysis was initially demonstrated in the diagnosis of ovarian cancer (65). In this study, exceptional results were seen using 5-20 specific key proteins identified, with a sensitivity and specificity of >95%, which is far superior to the sensitivities and specificities obtained with current serological cancer biomarkers. Subsequently, proteomic pattern analysis has been evaluated in a number of other cancer types, including breast, liver, and pancreatic cancers (66-68).

Two studies in breast cancer have investigated the correlation between SELDI-TOF mass spectrometry (MS) protein profiles of 105 tumour tissue lysates (69) and 27 breast cancer cell lines (26, 70). In both studies, patient subgroups identified by hierarchical clustering of SELDI-TOF MS protein profiles were analogous to the molecular breast cancer subtypes (69, 70). Of the several differentially expressed protein peaks detected, heat shock protein (Hsp) 27 and annexin V were identified as over-expressed in the luminal A type tumour tissue lysates (69), while S100-A9 (higher in basal) and a C-terminal truncated form of ubiquitin (higher in luminal) were found differentially expressed between the luminal-like and basal-like cell lines (70). Notably, subsequent IHC analysis of S100-A9 in tumour specimens of 547 early breast cancer patients confirmed its association with basal subtypes, as well as its value as an indicator of poor prognosis (70).

ii. Prognostic markers

In contrast to diagnostic studies, protein profiling studies aimed at discovering novel protein markers to prognosticate breast cancer are much more limited. One study (71) investigated the post-operative sera of 83 high-risk (mainly lymph node positive) breast cancer patients by SELDI-TOF MS and constructed a 40-protein signature that accurately predicted outcome in 83% of patients. The major components of this signature included haptoglobin alpha-1, complement component C3a, transferrin, and apolipoprotein A-I and C-I. These results should however be interpreted cautiously, as the number of proteins used for prognostication was rather high in comparison with the limited study population, indicating possible over-fitting of the data.

In another SELDI-TOF MS study performed in 60 breast cancer tissues, high levels of ubiquitin and/or low levels of ferritin light chain were found associated with a good prognosis (72). Although the results have not been confirmed by analysis of independent sample sets, ubiquitin has also been found differentially expressed in breast cancer subtypes by three other studies investigating tissue specimens (73) and cell lines (70, 74).
iii. Predictive markers

Several SELDI-TOF MS peaks (not structurally identified) were found indicative of treatment response in breast cancer cell lines to doxorubicin or paclitaxel (75). In addition, one study (76) found an increase of a 7.6kDa bovine transferrin fragment in serum-free conditioned medium of paclitaxel-resistant human breast cancer cell lines, corresponding to the increased expression of the transferrin receptor they observed in whole cell lysates. Although these results were not translated to the human in vivo setting, other studies have indeed reported an association between increased serum and cerebrospinal fluid transferrin levels and poor clinical outcome (71, 77). In one study, ubiquitin and S100-A6 were found to decrease in lysates of human breast cancer cell lines following chemotherapy-induced apoptosis (74); this coupled with the fact that aberrant expression of both proteins has also been reported in breast cancer tissue could make these two markers useful in predicting chemoresistant breast cancers (72, 73). In addition to these in vitro studies, in vivo studies have been performed as well (78, 79). In serum, both high molecular weight kininogen and apolipoprotein A-II were found to be significantly decreased in expression following docetaxel infusion in one particular patient with severe docetaxel side effects as compared to the other patients who tolerated the docetaxel infusion well. (79). The findings of this provocative study, if confirmed, suggest the potential of measuring protein biomarkers to predict adverse reaction to a drug.

B. Neoadjuvant chemotherapy and post-chemotherapy time point of biomarker analysis in breast cancer

Whilst the baseline pre-treatment time point is the commonest time point used in obtaining biomarkers to provide prognostic and/or predictive information, there are merits to using a post-chemotherapy time point biomarker, which may provide insights into biological effects of drugs and mechanisms of drug resistance. This however can only realistically occur in the setting of neoadjuvant or primary chemotherapy and in tumours from which serial tissue sampling can be safely obtained, such as in primary breast cancer. Neoadjuvant or primary chemotherapy in large primary breast cancers has been used with the purpose of reducing tumour volume and permitting less aggressive surgery (80). However, about 10-20% of patients do not benefit from this clinical approach (81, 82), and early identification of these patients could help avoid side effects from non-effective chemotherapy and unnecessary delay of definitive surgery.

1. Feasibility and significance of evaluating serial changes in protein expression post (neoadjuvant) chemotherapy in breast cancer

Almost a decade ago, one of the earlier studies (83) assessed the feasibility of obtaining serial core breast biopsies, and correlated rates of apoptosis, proliferation, and expression of related proteins at baseline, during, and after neoadjuvant single agent chemotherapy for locally advanced breast cancer with treatment response. The study recruited women with a histologically confirmed unresected T3 or T4 infiltrating carcinoma of the breast. The first 20 patients received three cycles of doxorubicin
90mg/m² followed by three cycles of paclitaxel 250mg/m², or the reverse. Nine women received four cycles of each (doxorubicin 60mg/m² and paclitaxel 175mg/m²). The end points studied included: clinical and pathological response, serial apoptotic [terminal deoxynucleotidyl transferase (Tdt)-mediated nick end labeling] and proliferation rates, and expression of ER, HER2, Bcl2, and p53 by IHC. Twelve patients (42%) had a clinical complete response (cCR), and 16 (55%) had a clinical partial response. Five women (17%) had pCR, 7 (24%) had microscopic residual disease, and 17 (58%) had macroscopic residual disease. Higher baseline apoptosis and proliferation were associated with a statistically significant improved pCR rate. In addition, among 14 evaluable patients, apoptosis increased in women who had a cCR to the first agent but not in women without a cCR. The study however did not show any serial changes in ER, Her2, Bcl2 or p53. The authors concluded that it was feasible to obtain serial core biopsies that are informative for studies of apoptosis and IHC in patients undergoing neoadjuvant chemotherapy.

2. Limitations of post-chemotherapy biomarker analysis

While feasible, for the most part, post-chemotherapy biomarker analysis is likely to be less well accepted by patients. This is because of all the accompanied logistical and patient discomfort issues with repeated biopsies. There would also be the issue of sampling error: biopsy and analysis of a chemotherapy-induced necrotic part of the tumour versus a still viable part or even chemo-resistant part of the tumour may reveal completely different profiles. It is also unclear at this point in time if any predictive biomarker for response obtained after treatment would be superior to standard clinical or radiological measurement of response. Having said that, many of the above issues also plague baseline biomarker analysis; a good example would be that of the recently recognized issue of Her2 heterogeneity in breast cancer (84).

C. Specimen sources for measuring changes in protein expression

1. Blood/ plasma/ serum

Since whole blood is considered to provide a dynamic reflection of physiological and pathological status, human plasma and serum represent the most extensively studied biological matrices in the quest for (breast) cancer biomarkers (85). Besides the usual circulatory proteins, it also contains specific tumour-secreted proteins, normal tissue- and plasma-proteins digested by tumour-secreted proteases, and proteins produced by local and distant responses to the tumour (86, 87). Several proteomic studies on plasma or serum utilizing MALDI-TOF MS and SELDI-TOF MS peaks have been reported to differentiate patients with breast cancer from those with benign breast disease and/or healthy controls (78, 88, 89).

Blood plasma is the liquid component of blood in which the blood cells in whole blood are normally suspended. It makes up about 55% of the total blood volume, and is the intravascular fluid part of extracellular fluid, comprising mostly water (93% by volume) and contains dissolved proteins, glucose, clotting factors, mineral ions, hormones and carbon dioxide. Blood plasma is prepared by centrifuging a tube of fresh blood
containing an anti-coagulant until the blood cells fall to the bottom of the tube. The blood plasma is then drawn off. In contradistinction, blood serum is blood plasma with the clotting factors removed by letting a collected tube of blood clot and the ensuing liquid portion aliquotted off. This would thus require less equipment than collecting blood plasma. Serum contains all proteins not used in blood clotting (coagulation) and all the electrolytes, antibodies, antigens, hormones and any exogenous substances. Plasma specimens may thus be analyzed for biomarkers related to the coagulation cascade, unlike serum specimens where the coagulation factors would have been consumed in the clotting process.

The commonest clinical use of blood instead of tissue biopsy to assess a tumour’s status serially in breast cancer would be the use of Ca15-3 (Section A5) as a surrogate for tumour response in metastatic breast cancer. The serial decrease in Ca15-3 in response to treatment is often congruent with the imaging findings of a response to chemotherapy even though it is based on expert panel (ASCO) recommendations rather than rigorous prospective data.

The most promising use of blood instead of tissue biopsy for measuring serial changes in protein expression would be in the area of Her2 oncoprotein. Other blood markers (e.g. osteopontin) showing serial changes of possible prognostic significance are discussed below. (Section E)

2. Tumour tissue

Tumour tissue can be collected fresh, ‘snap’ frozen in liquid nitrogen, or in formalin and then fixed in paraffin. The former is much more labour and logistics intensive while the latter has the potential problems of protein degradation from the fixation process. As it stands now most protein biomarker analysis are done on paraffin-fixed tissue due to the low cost and ease of transport. Fresh or fresh frozen tissue can be subject to MALDI/SELDI-TOF analysis but paraffin-fixed tissue can essentially only be used for IHC assessment.

3. Cerebrospinal fluid (CSF)

Besides blood, CSF has also been explored for cancer biomarkers (77, 90). CSF contains less total protein than serum and provides a low fluid-volume-to-organ ratio, thereby augmenting biomarker discovery (91). As collection of CSF by invasive lumbar puncture is not applicable to healthy controls, the studies thus far only have been for diagnosis of metastatic disease in the leptomeninges or for prognosis rather than for primary diagnostic purposes in breast cancer. In one study which aimed to search for markers indicative of leptomeningeal metastases, CSF samples of 106 breast cancer patients were digested with trypsin (77); the resulting peptides were then analysed by MALDI-TOF MS and a 164 peak classifier with 77% accuracy in determining leptomeningeal disease was constructed. The discriminative tryptic peptides were derived of several proteins (90), three of which (i.e. apolipoprotein A-I, haptoglobin and transferrin) have also been found to be associated with clinical outcome in serum (71).
4. Urine

Urine has also been looked at as a source of biomarkers for breast cancer due to its ease of collection. One study looked at matrix metalloproteinase-9 (MMP-9) and a disintegrin and metalloprotease 12 (ADAM12) and found that they could predict women who were at increased risk of developing breast cancer (92).

5. Nipple aspirate

Nipple aspirate and nipple ductal lavage have been investigated as a source of biomarkers; the rationale being that tumour cells could secrete proteins into the ducts. One study looked at nipple aspirate and ductal lavage specimens in patients with and without breast cancer, and found that elevated human neutrophil peptide in high risk cancer-free women, defined as those with estimated 5-year Gail risk of >1.6% or history of lobular carcinoma in situ, could predict early onset breast cancer better than current detection methods (93).

6. Tumour lysates

Tumour lysates are harvested from fresh or fresh frozen tumour samples. They are homogenized in a lysis buffer with protease inhibitors to prevent protein degradation. The sample is then centrifuged and the supernatant decanted to obtain the tumour lysate, which can be used for MALDI-TOF or SELDI-TOF, reverse phase protein array (RRPA; see below), and other high throughput proteomic analyses.

7. Circulating tumour cells (CTCs)

It has been appreciated in the past several years that tumour cells are shed from the primary tumour/metastases and circulate in the bloodstream. Intense research is ongoing to determine the utility of the detection of these cells in prognostication and prediction of therapy. One study showed that captured CTCs are amenable to biomarker analyses such as Her2 status, quantitative RT-PCR for breast cancer subtype markers, KRAS mutation detection and EGFR staining by immunofluorescence. The study was able to determine Her2 status by immunofluorescence and FISH in CTCs from metastatic breast cancer patients, although concordance with tumor Her2 status was only 89% (94).

D. Methods of measuring protein expression and its changes

1. Enzyme-Linked ImmunoSorbent Assay (ELISA)/ antibody microarray

In ELISA, an unknown amount of antigen is affixed to a surface, and then a specific antibody is applied over the surface so that it can bind to the antigen. This antibody is linked to an enzyme, and, in the final step, a substance containing the enzyme's substrate is added. The subsequent reaction produces a detectable signal, most commonly a color change in the substrate.

Performing an ELISA involves at least one antibody with specificity for a particular antigen. The sample with an unknown amount of antigen is immobilized on a solid support either non-specifically (via adsorption to the surface) or specifically (via capture by another antibody specific to the same antigen, in a "sandwich" ELISA). After the antigen is
immobilized, the detection antibody is added, forming a complex with the antigen. The detection antibody can be covalently linked to an enzyme, or can itself be detected by a secondary antibody that is linked to an enzyme through bioconjugation. Between each step, the plate is typically washed with a mild detergent solution to remove any proteins or antibodies that are not specifically bound. After the final wash step, the plate is developed by adding an enzymatic substrate to produce a visible signal, which indicates the quantity of antigen in the sample. The commonest medical use of ELISA is for assaying antigens/substances that can be found in the serum component of blood or bodily fluids (e.g., human immunodeficiency virus).

An antibody microarray is a specific form of ELISA-based protein microarray; a collection of antibodies are fixed on a solid surface such as glass, plastic or silicon chip, for the purpose of detecting antigens. The antibody microarray is often used for detecting protein expression from cell lysates and special biomarkers from serum or urine for diagnostic applications.

2. Tumour Immunohistochemistry (IHC)

Immunohistochemistry, which is the most practical method for assessing protein expression changes, not only provides a semi-quantitative assessment of protein abundance but also reveals cellular localization. Because no special processing of tissue samples is needed and labour intensive and expensive diagnostic techniques are avoided, IHC is perhaps the most readily adaptable technique to clinical practice. Inter-observer reproducibility of immunohistochemical scoring is sometimes problematic, although this can generally be resolved by re-evaluation and discussion to reach consensus. Analysis of protein expression using IHC has identified molecular subtypes in breast cancer that are similar to those derived from gene expression arrays (95). The most powerful use of IHC is that of a tissue microarray (TMA), which analyzes simultaneously a new protein marker or a group of ‘protein signature’ markers in hundreds to thousands of cylindrical fragments of clustered tumour samples collected from original paraffin blocks (96).

3. Proteomics: MALDI-TOF/ SELDI-TOF

Proteomic-pattern profiling is a recent approach to protein biomarker discovery. Given that mRNA information does not always accurately reflect the function of proteins which are the functional components within organisms, the use of proteomic patterns to enable tumour diagnosis or sub-classification seems more promising. The rationale is that proteins produced by cancer cells or their microenvironment may eventually enter the circulation and that the patterns of expression of these proteins could be assessed by mass spectrometry (MS) in combination with mathematical algorithms for diagnostic purposes. The search for novel protein biomarkers has for the longest time been dominated by two-dimensional gel electrophoresis (97), a significant disadvantage of which is its lack of real high-throughput capability. However, recent advances in analytical technologies, such as protein microarrays and mass spectrometry, have enabled large-scale proteomic analyses (98).

Technologies such as differential in-gel electrophoresis, two-dimensional polyacrylamide gel electrophoresis and multidimensional protein identification technology, can be used for
high-throughput protein profiling. Due to their relative ease of sample preparation, high analytical sensitivity and speed of data acquisition, two MS-based technologies in particular, matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) MS (99) and its variant surface-enhanced laser desorption/ionisation (SELDI-) TOF MS (100) have been widely deployed for cancer biomarker discovery (101). In both platforms, biological samples (e.g., serum, tissue lysate) are co-crystallised with an energy absorbing matrix on a sample probe surface. Subsequent irradiation with brief laser pulses sublimes and ionises the proteins out of their crystalline matrix, after which an electric field migrates the charged proteins to the time-of-flight mass analyser. The proteins are then separated based on their mass, as the time to detector impact (TOF) is proportional to the protein mass per charge. The two platforms differ in their sample probe surfaces. In MALDI, the probe surface merely presents the sample to the mass spectrometer, warranting off-line sample fractionation and clean-up to produce usable MS signals. In contrast, the probe surfaces utilised by SELDI are comprised of various chromatographic surfaces, enabling their active role in sample fractionation.

The technology that has received considerable attention involves the use of a minute amount of biological sample added to a ‘protein-chip’, which is subsequently analyzed by MALDI or SELDI-TOF-MS to generate a proteomic signature (102). These patterns reflect part of the tissue or body fluid proteome, but without knowledge of the actual identity of the proteins. In addition to the issue of protein identification, there are also problems related to the other aspects of this technology. Validation and the consistency of bioinformatics analysis is of great importance to ascertain reproducibility and prevent systematic bias and overfitting of data. This is highlighted by a study (103), in which the potential markers for breast cancer and lymph node status reported by two studies (104, 105) could not be confirmed following analysis of an independent sample set. The shortcomings also include bias from artefacts related to the clinical sample collection and storage, the inherent qualitative control issues of mass spectrometric analysis, failure to identify well-established cancer biomarkers, bias when identifying high-abundance molecules within the serum, and disagreement between peaks generated by different laboratories (106, 107). Another limitation concerns possible bioinformatic artifacts; one study (108) showed that even signals that are detected that are actually a result of ‘noise’ can also achieve a high level of discrimination between patients with cancer and those without, further highlighting the lack of specificity of some of the signals detected. As such, despite a significant period since the first report of this technology, no independent validation studies have been published, and no product has yet reached the clinic.

4. Isobaric Tags for Relative and Absolute Quantification (iTRAQ)

iTRAQ is based on the labeling of the N-terminus and side chain amines of peptides from digested proteins with tags of varying mass. One such method commercialized by Applied Biosystems® is called iTRAQ and uses four amine specific (4-plex) isobaric reagents to label the primary amines of peptides from four different biological samples. In recent times, an eight amine specific set of reagents (8-plex) has also been available. These samples are then
pooled and usually fractionated by nano liquid chromatography and analyzed by tandem mass spectrometry. A database search is then performed using the fragmentation data to identify the labeled peptides and hence the corresponding proteins. The fragmentation of the attached tag generates a low molecular mass reporter ion that can be used to relatively quantify the peptides and the proteins from which they originated. iTRAQ reagents therefore allow simultaneous identification and quantification of proteins in four or eight different samples using tandem mass spectrometry.

5. Reverse Phase Protein Array (RPPA)

The principle of RPPA involves the spotting of patient samples in an array format onto a nitrocellulose support. Hundreds of patient specimens can be spotted onto the same array, allowing a large number of samples to be compared simultaneously under identical conditions. Each array is incubated with one particular antibody, and signal intensity proportional to the amount of analyte in the sample spot is generated. Signal detection is commonly performed by fluorescence, chemiluminescence or colorimetric methods, and the results are quantified by scanning and analyzed by software such as P-SCAN® and ProteinScan®.

RPPA is possibly a useful tool to identify and validate proteins and phospho-proteins in cancer (109, 110). The aim of one such study was to determine whether functional proteomics using RPPA improves breast cancer classification and prognostication and also whether it can predict pCR in patients receiving neoadjuvant taxane and anthracycline-taxane-based chemotherapy. Six breast cancer subgroups were identified by a 10-protein biomarker panel in the 712 tumour training set, that were associated with different recurrence-free survival. A prognosis score constructed using the 10 protein-signature (ER, PR, Bcl2, GATA3, CCNB1, CCNE1, EGFR, Her2, Her2p1248 and EIG121) in the training set was associated with relapse free survival in both the training and test sets. In addition, there was a significant association between the prognostic score and likelihood of pCR to neoadjuvant chemotherapy in yet another independent sample set.

E. Serial changes in expression of various established protein biomarkers in response to neoadjuvant chemotherapy

1. ER/PR

The most obvious candidate biomarker for which one would expect extensive data on with regards to serial changes post chemotherapy would be ER and PR. Much as it would have been hoped that the data on this would be consistent, the fact is that various trials showed a spectrum of findings. In a relatively recent meta-analysis (111), it was found that discordance in ER and PR between core needle biopsy and subsequent resection material was more evident in the patients treated with neoadjuvant chemotherapy (around 15%) than the reported discordance in patients not treated with neoadjuvant chemotherapy (around 2%). Although the studies reviewed were quite heterogeneous with respect to study methodology, design and outcome measures, the discordances could only partly be explained by the study design confounders and are instead more likely due to the direct effect of the chemotherapy. A change of hormone receptor status in up to a third of patients
after neoadjuvant chemotherapy was reported in several studies (112-114). In general, studies that reported a good concordance of hormone receptor status before and after neoadjuvant chemotherapy had a relatively smaller number of patients compared to the studies that found significant changes. The small sample size might have prevented these trials from showing statistical significance. Intriguingly, PR was found to be more discordant compared to ER.

Various postulations regarding the mechanisms resulting in a change in the receptor status caused by chemotherapy have been put forth. The targeting of chemosensitive tumour cells leaves resistant tumour cells behind, which may have different ER status (positive or negative) from the sensitive tumour cells that were eradicated. Lower circulating levels of estrogens caused by ovarian insufficiency during or after chemotherapy in premenopausal women (115) might cause downregulation of the estrogen and/or progesterone receptor of the tumour leading to estrogen-independent growth.

In spite of the likely true observed phenomenon of serial changes in ER/PR status post chemotherapy, however, little is known about the predictive or prognostic value of a changed receptor status. A few investigators tried to correlate changes to treatment response, but discordant conclusions were drawn (116, 117). A positive switch of the hormone receptor status could be an indicator for a better outcome and indeed was significantly correlated with better disease-free and overall survival in patients who were treated with adjuvant endocrine therapy compared to those with a positive switch who were not (118). In clinical practice, the likely utility of repeating the ER/PR status post chemotherapy is to evaluate if there is any hormone receptor positivity necessitating endocrine therapy.

2. Tumor Her2

In one study (119), the authors evaluated the correlation among patients’ characteristics, immunohistochemical expression of hormonal receptors (ER and PR), p53, p21 and Her2 protein expression and the clinical and pathological response to a neoadjuvant combination of docetaxel and epirubicin chemotherapy. There was a reduction in p53 protein expression, as well as in p21 protein expression after neoadjuvant chemotherapy. However, neoadjuvant taxane and anthracycline did not change Her2 expression in patients with locally advanced breast carcinoma. The tissue Her2 stable phenotype observed in this study during neoadjuvant chemotherapy has also been reported by other groups (112, 117, 120). Recent studies have focused on the role of tumour Her2 as a predictive factor of response to neoadjuvant chemotherapy (112) but have failed to observe a correlation between tumour baseline Her2 positivity and the clinical or pathological response to neoadjuvant chemotherapy. This is in contradistinction to serum Her2 which will be detailed below.

3. Serum Her2

There have been several studies showing serial changes in serum Her2 (sHer2) following neoadjuvant chemotherapy. One study within a large clinical trial (121) sought to use serum markers to optimize treatment strategies in breast cancer. The authors investigated serum
Her2 levels (sHer2) in 175 breast cancer patients participating in the GeparQuattro trial. This study incorporated neoadjuvant chemotherapy approaches in Her2-positive and negative patients (epirubicin/cyclophosphamide prior to randomization to either docetaxel alone, docetaxel in combination with capecitabine or docetaxel followed by capecitabine) and the addition of trastuzumab treatment for patients with Her2-positive tumours. sHer2 levels were measured by ELISA before and after initiation of neoadjuvant chemotherapy in 90 Her2 positive and 85 Her2 negative patients. Median pre-chemotherapy sHer2 levels were higher in patients with positive Her2 status of the primary tumour than in patients with negative tumor Her2 status (14.9ng/ml versus 7.7ng/ml). A pre-chemotherapy sHer2 cut-off level of 10ng/ml had the best sensitivity and specificity in differentiating between Her2 positive and Her2 negative tumours. In Her2 positive patients, the authors found a significant positive association between pCR and elevated baseline sHer2 levels (above 15ng/ml) and a more than 20% decrease of sHer2 levels during neoadjuvant chemotherapy, which was of borderline significance in multivariate analysis (odds ratio=3.29). In Her2 negative patients, the authors observed no association between sHer2 levels and pCR. The authors thus hypothesized that monitoring sHer2 levels in the presence of anti-Her2 treatment might be an adjunct to clinical evaluation during neoadjuvant chemotherapy in Her2 positive disease.

Two smaller published reports investigated the correlation between treatment-induced changes in sHer2 and pathologic complete response from neoadjuvant chemotherapy plus trastuzumab. One study evaluated sHer2 levels in a trastuzumab-based neoadjuvant setting in 16 patients. In this small group of patients, the authors could show that a decrease of sHer2 levels was associated with response to therapy (122). In the other study which monitored sHer levels serially over 6 months, 39 patients were treated with neoadjuvant chemotherapy including 29 patients who received a trastuzumab combination. A 9% decrease in sHer2 levels from week 3 to week 6 after initiation of therapy (but not earlier or later) was predictive of pCR (123). This study also illustrated that time dependency of post chemotherapy biomarkers could be an important issue. In this study, the mean sHer2 baseline values were not different between the pCR group and the group with residual disease suggesting that post-chemotherapy evaluation could be superior to baseline evaluation.

In contrast, Quaranta et al. could not find a correlation between serum positivity for Her2 (using a cutoff of 10ng/ml as per the GeparQuattro trial) and tissue positivity for Her2 levels in an unselected patient group of 108 patients (124). In addition, no clear relationship was found between baseline sHer2 levels and tumour response to trastuzumab-based treatment in a recently published meta-analysis (125). In an abstract of a small study presented at ASCO 2011, Lee et al investigated the correlation between the response of advanced breast cancers to neoadjuvant chemotherapy and the change of serum Her2 and short-term disease free survival. Twenty-two locally advanced Her2 IHC 3+ or FISH amplified breast cancer patients were treated with neoadjuvant doxorubicin or trastuzumab. Serum Her2 levels were measured by chemiluminescence immunoassay before and after neoadjuvant chemotherapy. The cutoff value was 10.2mg/ml which is similar to that of the GeparQuattro
Serial Changes in Expression of Proteins in Response to Neoadjuvant Chemotherapy in Breast Cancer

Mean serum Her2 before chemotherapy was 15.8± 1.6ng/ml, and that after chemotherapy was 10.6 ± 0.38ng/ml. The change of serum Her2 in the CR group was higher than that in the PR group (13.26 ± 14.1ng/ml and 2.74±3.2ng/ml respectively). However, at a mean follow-up of 41 months, the change of serum Her2 before and after chemotherapy was not correlated with disease recurrence or with disease free interval.

Several groups have postulated that monitoring changes in sHer2 levels after a specific time-period after trastuzumab treatment might be valuable for identifying a patient population that might benefit from additional treatment regimens with other Her2 targeted therapies and certainly merits confirmation in further large prospective trials (126).

F. Serial changes in expression of various candidate protein biomarkers in response to neoadjuvant chemotherapy

1. COX-2

Cyclooxygenases (COX-1 and COX-2) are rate-limiting enzymes in the formation of prostaglandins from arachidonic acid. COX-1 is considered to be constitutively expressed while COX-2 is highly inducible by various factors and is associated with tumourigenesis (127-129). Several studies have shown the unfavourable prognostic significance of COX-2 expression in breast cancer (130, 131). A few retrospective breast cancer studies have also suggested that tumour expression of COX-2 may be associated with more aggressive breast cancer phenotypes, poorer response to chemotherapy and inferior survival (130, 132). One study compared serial tumour samples from individual breast cancer patients before and after exposure to sequential cycles of doxorubicin and docetaxel and examined changes in tumour expression of COX-2 by IHC. The study also correlated any significant changes in biomarker expression with tumour clinical response and progression-free survival. There was a statistically significant progressive downward trend in COX-2 expression with increasing cycles of chemotherapy for the entire cohort. Subgroup analysis found that this decrease in COX-2 expression to be predominant in clinical responders but not in non-responders. COX-2-positive tumours at baseline showed a statistically significant reduction in COX-2 expression with chemotherapy. This downward trend was most marked between the third and sixth cycle of chemotherapy rather than after one cycle of chemotherapy suggesting that this change did not occur early during chemotherapy. Tumours that were COX-2 positive both at baseline and after treatment had the worst outcome, while those that were COX-2 negative both at baseline and after treatment had the best outcome with a median progression free survival (PFS) of 25 versus 47 months. Another significant finding is related to ER status and COX-2 overexpression. For ER-positive and COX-2-positive tumours at baseline, a change to COX-2 negativity resulted in a statistically significant improvement in PFS compared with tumours that remained COX-2 positive (52 versus 27 months). As for ER-negative and COX-2-positive tumours at baseline, the PFS is generally poor regardless of whether the tumour remained COX-2 positive or became negative after chemotherapy. The study also showed that COX-2-positive tumours at baseline correlated with more advanced tumour size, presence of metastases and inferior PFS as compared with COX-2-negative tumours. That neoadjuvant chemotherapy resulted in a reduction in COX-2 expression in breast tumours is consistent with findings previously observed in breast
cancer cell lines after chemotherapy (133). Furthermore, this reduction in COX-2 expression was seen mainly in clinical responders, a phenomenon that is also documented in other cancers (134). Another interesting finding in this study was that in patients with ER-positive and COX-2-positive tumours at baseline, post-chemotherapy COX-2 positivity had a significant negative influence on PFS, suggesting that COX-2 could play an important role in hormone-dependent breast cancers.

The interest in COX-2 expression in cancer arises from the fact that this over-expression occurs in many human malignancies including colon and lung cancer (131, 135, 136), and the possibility of using widely available COX-2 inhibitors, e.g., nonsteroidal anti-inflammatory drugs, together with conventional anticancer therapy to enhance treatment efficacy. This is based on the premise that many of the COX-2-regulated genes that contribute to tumour progression may also be determinants of tumour sensitivity to treatment (137). Although the potential chemopreventive properties of selective COX-2 inhibitors are being actively investigated, little is known about the utility of these agents in the treatment of cancer. However, there is emerging data from breast, pancreatic and lung cancer studies showing potential benefit of combining COX-2 inhibitors with chemotherapy (137-139). In particular, a recent study involving patients with heavily pretreated breast cancer showed that the combination of chemotherapy and a selective COX-2 inhibitor resulted in a statistically significant doubling of time to progression for COX-2-positive tumours compared with COX-2 negative ones. At this point in time, there have not been confirmatory studies on COX2 serial changes post chemotherapy in breast cancer but this appears to be an area worth exploring. In particular, post-chemotherapy COX-2 expression may serve as a biomarker for COX-2 inhibition as a therapeutic strategy that warrants evaluation.

2. **Topo2α**

A seminal study (140) analyzed the value of Topo2α in predicting clinical response to anthracycline-based neoadjuvant chemotherapy in breast cancers and its potential changes after chemotherapy. The study also looked at p53 and Her2 the latter being commonly coexpressed with Topo2α. Forty-one patients with primary breast cancer and treated with neoadjuvant anthracycline-based chemotherapy were included in the study. Topo2α, Her2 and p53 expression were measured by IHC in pre- and post-chemotherapy tumour specimens and the results were correlated with clinical response. Topo2α was overexpressed in 16 of 41 (31%) tumours before treatment, and this baseline overexpression was significantly associated with clinical response. Of note, Topo2α overexpression, but not Her2 or p53, was lost in specimens after chemotherapy, although this change did not correlate with clinical response. The observed link between baseline Topo2α expression and clinical response to neoadjuvant anthracycline-based chemotherapy, together with its loss after chemotherapy, suggests that Topo2α deserves further testing in a prospective setting as a predictive marker for chemotherapy response.

3. **MGMT**

O6-Methylguanine-DNA methyltransferase (MGMT) rapidly reverses alkylation (including methylation) at the O6 position of guanine by transferring the alkyl-group to the active site
of the enzyme, constituted by a cysteine. An inactivated MGMT gene allows accumulation of O6-alkylguanine that is the most cytotoxic lesion of alkylating agents, which subsequent to incorrect pairing with thymidine triggers mismatch repair, thereby inducing DNA damage and eventually cell death. There has been one recent trial published by a Japanese group showing the possible utility of MGMT in breast cancer (141). The study evaluated thirty-two basal like breast cancer patients receiving neoadjuvant chemotherapy with an anthracycline and taxane-based regimen. The immunoreactivities of MGMT, MLH1, MSH2 and BRCA1 before and after neoadjuvant chemotherapy were evaluated. pCR was achieved in 10 of 32 cases (31%), and cancer-related and disease-free survival rates were significantly higher in the pCR group than in the non-pCR group. In biopsy samples before neoadjuvant chemotherapy, attenuated expression of MGMT, MLH1, MSH2 and BRCA1 was observed in 12/32 (38%), 0/32 (0%), 5/32 (16%) and 28/32 (88%) cases, respectively. On evaluation of predictors of pCR, including patient characteristics (age, menopausal status, clinical and pathological stages) and immunohistochemical patterns, pre-chemotherapy reduced expression of MGMT was found to be the only factor significantly predictive of pCR. Paired biopsy samples before neoadjuvant chemotherapy and surgical tumour material after neoadjuvant chemotherapy were available for 19 cases of non-pCR. In these 19 cases, decrease in expression of MGMT during neoadjuvant chemotherapy was more frequently observed in those with tumour shrinkage (i.e. > 60%) as compared to those with no decrease, although the difference was not statistically significant possibly due to the small sample size. If these results can be validated, baseline MGMT expression may be used as a predictor of pCR to neoadjuvant chemotherapy, while decrease in MGMT expression with chemotherapy may have additional predictive value for treatment response.

4. Heat shock protein 70

Recently one group has shown using proteomic analysis (2-dimensional gel electrophoresis and mass spectrometry) of fourteen matched pairs of ER positive tumour tissues before and after neoadjuvant treatment with an aromatase inhibitor (AI) that ten proteins were differentially expressed before and after AI treatment. Among the identified proteins, treatment-induced reduction in heat shock protein 70 (Hsp-70) expression was the most significantly correlated with both clinical and pathological responses (142). This downregulation of Hsp-70 with chemotherapy was subsequently confirmed by IHC. These findings suggest that Hsp-70 may represent a potential novel predictive marker to endocrine therapy response (143).

5. Osteopontin

Osteopontin has been reported to be a malignancy-associated protein measurable in tumour tissue and blood. In a prospective clinical study that measured serial plasma osteopontin levels in women with metastatic breast cancer throughout the course of their disease, serial elevation of osteopontin was found to be prognostic (144).

One hundred fifty-eight women with newly diagnosed metastatic breast cancer were enrolled in the study. Plasma osteopontin was measured using an ELISA assay, at baseline
and every 3 to 12 weeks during and after therapy until death. Multivariate time-dependent survival analyses were conducted using models that right censored patient outcomes 3, 6, and 12 months after the last known osteopontin measurement. Osteopontin was measured in 1,378 samples (median, 9 per patient). Ninety-nine patients had elevated baseline osteopontin (median, 177 ng/ml; range, 1-2,648 ng/ml). In univariate analysis, elevated baseline osteopontin was associated with shorter survival (p = 0.02). In a multivariate model incorporating standard prognostic factors, baseline osteopontin was only marginally significantly associated with survival duration (relative risk, 1.001; p = 0.038). However, in a multivariate model incorporating standard prognostic factors and changes in sequential osteopontin levels, an osteopontin increase of >250 ng/mL at any time was the variable with the most prognostic value for poor survival (relative risk, 3.26; p=0.0003).

This study is a further proof of concept that serial changes (and not just a baseline value) in a blood biomarker (and not just a tumour biomarker) can be prognostic in breast cancer.

6. Serial changes in protein profiles identified by high-throughput assays

The study of serial changes in protein profiles in tumour and/or serum induced by chemotherapy has been assessed in a high throughput fashion. In a proteomics study, comparison of protein profiles using MALDI-TOF analysis of sera acquired before and after preoperative chemotherapy for breast cancer was performed (145). The study analyzed pre- and post-chemotherapy protein profiles of sera from 39 Her2-positive breast cancer patients who received 6 months of preoperative chemotherapy using liquid chromatography-MALDI-TOF/MS technology, and detected qualitative and quantitative differences in pairwise comparison of pre- and post chemotherapy samples that were different in the 21 patients who achieved pathological complete response compared with the 18 patients with residual disease. 2349 and 3152 peaks were identified as differentially expressed in the pre-chemotherapy samples of the responders and non-responders respectively. Comparison of paired pre- and post-chemotherapy samples identified 34 (32 decreased, 2 increased) and 304 peaks (157 decreased, 147 increased) that significantly changed after treatment in responders and non-responders, respectively. The top 11 most significantly altered peptide peaks with the greatest change in intensity were also identified. These peaks matched eight different known proteins in an NCBI nr database search by MASCOT® software, including α-2-macroglobulin, complement 3, hemopexin, and serum amyloid P in the responder group and chains C and A of apolipoprotein A-I, hemopexin precursor, complement C, and amyloid P component in the non-responding group. All proteins decreased after therapy, except chain C apolipoprotein A and hemopexin precursor that increased. These results suggest that changes in serum protein levels occur in response to chemotherapy and these changes to a certain extent appear different in patients who are highly sensitive to chemotherapy compared with those who are more resistant.

7. Negative studies of post chemotherapy changes in biomarkers

Publication bias would tend to result in under-reporting of negative findings of changes in biomarkers post-chemotherapy. In the few negative trials published, some of them conflict with the findings of the other trials mentioned above. For example, in a study of 97 patients
who received neoadjuvant anthracycline-based chemotherapy (146), the authors failed to find any post-chemotherapy change in ER, Her2, p53, Ki67 or Bcl2 as assessed by IHC. This is in contradistinction to the study by Dawson et al (64) who found changes in Bcl2. One possible reason for the failure to find a change in the biomarkers post-chemotherapy is the small sample size in this trial.

G. The future of protein biomarkers in breast cancer

Studies of protein biomarkers in breast cancer still rely heavily on IHC and with a possible emerging role of ELISA. Newer technologies on the horizon could facilitate the discovery of novel biomarkers in a high throughput fashion, and there are a few interesting developments attempting to push the frontier in proteomics.

1. SILAC

SILAC (stable isotope labeling by/with amino acids in cell culture) is a MS-based methodology (used for quantitative proteomics) that detects differences in protein abundance using non-radioactive labeling. SILAC has emerged as a very powerful method to study cell signaling, post translation modifications, protein-protein interaction and regulation of gene expression.

2. Peptidomics

The low-molecular-weight plasma (serum proteome) has been the focus of recent attempts to find new biomarkers (147). Peptides are critical for many physiological processes, such as blood glucose (insulin) regulation. It has been suggested that “the low molecular-weight region of the blood proteome contains precious diagnostic information (148). The low-molecular-weight serum proteome has been characterized by ultrafiltration, enzymatic digestion, and liquid chromatography coupled to tandem mass spectrometry (149, 150), or via a top-down proteomics approach (whereby the intact peptide is distinguished directly by its fragment ions) (151) or by means of pattern profiling (152). Informative diagnostic peptides that are generated after proteolysis of high-abundance proteins by the coagulation and complement enzymatic cascades can be identified by mass spectrometry. These proteomic patterns were claimed to distinguish not only healthy controls from patients with cancer (153) but also between various types of cancer (152). However, one major concern is that these peptides present in the serum are derived from a small number of highly abundant proteins. One study showed that peptides in serum are affected by collection conditions. Improper collection could give rise to artefacts and serum is not ideal for proteomic experiments as it contains substantial endoproteolytic and exoproteolytic enzymatic activity (154). These findings raise concerns regarding peptidomics data generated by profiling technologies, with some investigators suggesting that peptidomic profiling might represent nothing more than peptides cleaved during coagulation or functions inherent to plasma or serum, including immune modulation, inflammatory response and protease inhibition (155). In addition, many of the issues associated with mass-spectrometry- based protein profiling technologies also apply to peptidomics. Thus, while this technology looks promising, more confirmatory data is required and awaited.
3. Cancer-biomarker-family approach

The basis for the ‘cancer biomarker family’ approach is that if a member of a protein family is already an established biomarker, then other members of that family might also be candidate cancer biomarkers. As an example, Prostate Specific Antigen (PSA) is a member of the human tissue kallikrein family. Kallikreins are secreted enzymes with trypsin-like or chymotrypsin-like serine protease activity. This enzyme family consists of 15 genes clustered in tandem on chromosome 19q13.4.63. PSA (KLK3) and KLK2 currently have important clinical applications as prostate cancer biomarkers (156). Other members of the human kallikrein family have been implicated in the process of carcinogenesis and are being investigated as biomarkers for diagnosis and prognosis. For example, KLK6 has been studied as a novel biomarker for ovarian cancer (157), and it was found that elevated serum levels of KLK6 was associated with late-stage tumour, high grade and serous histology and chemo-resistance. Similarly, KLK3, KLK5 and KLK14 have been shown to be increased in the serum of patients with breast cancer, thereby potentially serving as diagnostic markers. The fact that these proteins are serine proteases could implicate them in tumour progression through extracellular matrix degradation.

4. Secreted protein approach

Examination of tissues or biological fluids near to the tumour site of origin could facilitate identification of candidate biomarker molecules. The mounting evidence that tumour growth is dependent on the malignant potential of the tumour cells as well as on the microenvironment surrounding the tumour (e.g., stroma, inflammatory cells, etc) further supports this approach (158). A number of technologies can be employed for analysis of these samples, but for systematic characterization of proteins in complex mixtures, mass spectrometry is the preferred technology. In the case of breast cancer, breast tissue, nipple aspirate fluid, breast cyst fluid and tumour interstitial fluid can all be explored. The tumour interstitial fluid that perfuses the tumour microenvironment in invasive ductal carcinomas of the breast has been examined by proteomic approaches (159). Over 250 proteins were identified, many of which were relevant to processes such as cell proliferation and invasion. The identification of secreted proteins in tissues or other biological fluids does not necessarily imply that the proteins will be detectable in the sera of cancer patients though, as this will depend on the stability of the protein, its clearance, its association with other serum proteins and the extent of post-translational modifications.

20–25% of all proteins are secreted and/or undergo aberrant secretion of membrane-bound proteins that have a secretable/cleavable extracellular domain. Alterations in the signal peptide of proteins as a result of single nucleotide polymorphisms can result in unusual secretion patterns (160). Moreover, elevation of molecules in biological fluids can result from a change in the polarity of cancer cells, which can lead to the release of cancer-associated glycoproteins into the circulation. Increased expression of proteases that cleave the extracellular domain portion of membrane proteins can also cause increased circulating levels. One currently used secreted marker in breast cancer is that of CA15-3; it being a soluble form of MUC1, which is an antigenic focus on breast cancer. It is hoped that in the
years to come, more of such secreted proteins can be discovered so as to facilitate breast cancer monitoring.

2. Conclusion

Biomarkers guide physicians in counseling patients with regard to their prognosis and also provide information to physicians with respect to the optimal treatment for a particular patient/patient group. The latter scenario (predictive biomarkers) is an area in which the natural history of the patient can be affected positively and thus are deemed to be more important.

The baseline status of the predictive protein biomarkers ER, PR and Her2 are the most important in current breast cancer management. While post chemotherapy protein biomarkers in breast cancer is not currently used in routine clinical practice, evaluation of serial changes in expression of proteins in response to neoadjuvant chemotherapy has been shown to be feasible and in some studies shown to be a better biomarker than the baseline biomarker. It is achievable using an ELISA or IHC platform, which are technologies that are readily available in almost all clinical practices. However, the logistical challenges of obtaining serial tumour samples, reproducibility of the expression profile and patients’ compliance could be major factors that may limit widespread application of studying serial changes of protein expression in tumor routinely. There is also the issue of intra and inter-individual variability when examining for the presence of prognostic and/or predictive serial changes in any putative biomarkers. In fact such an issue already exists with regard to Her2 testing by FISH; the clinical significance of genetic heterogeneity (i.e. in certain tumours, only a fraction of the cells are positive for the oncprotein) of Her2 is still being investigated (84, 161).

Although newer technologies (e.g., MALDI-TOF, etc.) are promising, the inability to identify all ‘peaks’ and the reproducibility issues at this point in time are major limitations. However, development of methods that allows rapid characterization of identified protein peaks holds promise for more widespread use of these technologies in the near future.

Author details

Daniel Chan and Soo-Chin Lee
National University Cancer Institute Singapore

3. References


