ABSTRACT • Gene expression profiling has been increasingly used to determine new cancer markers. This technology holds major promises for improving the management of patients with breast cancer in which traditional clinicopathologic parameters do not account for all the heterogeneity of this disease and its distinct prognostic groups. Gene expression profiling has resulted in new classification of breast cancer and new assays are being developed and commercialized as prognostic and predictive tests.

However, the use of these tests in a clinical setting presents many issues. The accuracy of this new technology is often overestimated and its limitations should be addressed. Although early results are promising, further validation and well designed clinical trials are required before incorporating these tests in routine clinical practice.

INTRODUCTION

Breast cancer is the most common cancer in women. Current prognosis and management is based on clinical and histological parameters including TNM and tumor grade [1-2]. Despite their utility, these parameters fail to account for different clinical courses and outcomes in patients presenting with tumors showing similar features. Moreover, many women diagnosed with breast cancer get treatments from which they will not benefit and which may result in unnecessary toxicity and side effects.

Hundreds of markers have been tested in breast cancer in search for prognostic and predictive factors but despite years of research, only estrogen receptors, progesterone receptors, and HER2 have proven their utility in the routine clinical setting.

Technologic advances in the field of molecular biology have resulted in the ability to examine the expression of several genes simultaneously resulting in the recognition of patterns of gene expression. Current studies are showing that these patterns may serve as powerful tumor markers.

GENE EXPRESSION PROFILING. WHAT IS IT?

Gene-expression profiling is the simultaneous quantification of the expression of multiple genes. In other terms, it is the measurement of the activity of several genes in breast cancer cells. Such measurements are known as gene expression profiles. Several methods have been developed to look at the activity of hundreds or thousands of genes simultaneously. The most commonly used method is what we call “microarray analysis” or “gene expression array”. Another common method is reverse transcriptase-polymerase chain reaction (RT-PCR).

Detailed description of these methods is beyond the scope of this review. In brief, RT-PCR is the amplification detection and quantification of mRNA of a set of predetermined genes in a given sample. RT-PCR can be performed on fresh-frozen and formalin-fixed-paraffin-embedded tissues (FFPE).

Microarray analysis requires fresh-frozen tissues. It is performed by placing a segment of DNA representing a single gene onto a spot adhered to a solid medium, such as a glass slide or “chip”. Using this method, one can place hundreds or even thousands of DNA spots, each representing a single gene, onto a slide or chip, therefore creating an array of genes [3]. Next, RNA is extracted from the cells under study (such as breast cancer cells), amplified, labeled with a fluorescent dye, and then placed on the chip for hybridization with the complementary
DNA sequence present within the spots on the chip. In most cases a control RNA sample from normal cells (such as leukocytes) is placed on the same chip but labeled with a different fluorescent dye. The chip is then placed in a scanner and relative levels of fluorescence, from each spot, representing increased or decreased levels of RNA messages compared to the control, are collected and analyzed using special software and statistical techniques [3-4]. In most cases, only a minority of genes are found to be over or under-expressed resulting in what we call gene expression patterns allowing the recognition of tumors that share so-called “signatures.”

APPLICATIONS OF GENE EXPRESSION PROFILING IN BREAST CANCER

Gene expression profiling has been extensively used in breast cancer. It has resulted in refining our understanding, classification and grading of breast cancer as well as in the development of new prognostic and predictive tests.

Molecular classification of breast cancer

In their seminal paper Perou et al. [4] using microarray analysis identified distinct gene expression signatures, or “molecular portraits” when normal breast tissue was compared with breast cancer cells. The initial classification included basal-like, HER2 positive, luminal, normal, and categories of invasive breast cancer. Refining their earlier observation, these investigators reported that breast cancer could be grouped into six categories, which they designated basal-like; HER2-positive; normal; and luminal types A, B and C. Studies have shown that these classes are prognostic [5-7].

– **Tumors of the luminal groups** are ER positive and two thirds of them are of low or intermediate histological grade. Luminal A type breast cancers have the highest expression of estrogen and progesterone hormone receptors and have a better outcome than the other categories.

– **Basal-like cancers** are ER-negative and HER2-negative and most are high grade. The basal-like group has been associated with the so-called “triple-negative” breast cancer phenotype (ER negative, PR negative, and HER2 negative), but the basal-like genotype can be divided into multiple additional subgroups [8-9].

– **Basal-like and HER2-positive/ER-negative subtypes** are more sensitive to anthracycline-based neoadjuvant chemotherapy than are the luminal breast cancers [10]. They are also associated with the highest rates of pathologic complete response to neoadjuvant multiagent chemotherapy [11].

This molecular-based classification is not faultless and it appears that conventional parameters such as histological grade and hormonal receptors will complement the molecular classification rather than being replaced by it. Indeed, there are subtypes of luminal tumors that have poor outcomes despite being ER-positive and these consistently share the histopathological feature of being higher grade [6, 12-13]. Moreover, a recent statistical meta-analysis of previous studies concluded that only three subtypes are consistently identifiable across datasets: HER2 positive, ER positive/HER2 negative, and ER-negative/HER2 negative [14-15].

Genomic grading

Sotiriou et al. identified genes that were capable of reclassifying the traditional three histological grades of breast cancer into only two distinct molecular grades: High-Genomic Grade Index (GGI) and Low-GGI [16]. In that study, high and low GGI were strongly associated with histologic grades 1 and 3, but patients with histologic grade 2 tumors appeared to be similar to either the grade 1 or grade 3 tumors in their molecular profiles. Moreover, patients with histologic grade 2 and High-GGI had a higher risk for recurrence than patients with grade 2 and Low-GGI. These results suggest that histologic grade 2 is actually a mixture of grade 1 and grade 3 tumors. Loi et al. used GGI to define two ER-positive molecular subgroups of breast with a statistically distinct clinical outcome in both untreated and tamoxifen-treated patients [17].

It seems that the discrepancy between histologic grade and GGI can be accounted for by the instability of the mitotic figures component in the “Bloom-Richardson score” histologic grading system. Interestingly, van Diest et al. [18] reported that the tumor proliferative index assessed by Ki-67 immunostain could serve as a surrogate for mitotic figures counting, leading to better correlation between histologic and genomic grading.

Gene expression profiling prognostic and predictive tests

Gene expression profiling allowed the development of several prognostic and predictive tests for breast cancer. Currently the best knows tests are the Mammaprint and the Oncotype DX.

**Mammaprint** • This is the first commercially available, FDA approved, microarray-based test. It is currently designed as a prognostic assay for women under the age of 61 with lymph node negative breast cancer. It is performed on either fresh-frozen tumor samples or tumors preserved into an RNA preservative solution. The test results indicate either a high risk or low risk of disease recurrence. This test was developed in the Netherlands Cancer Institute. Using DNA microarray analysis on archived frozen tissue from 117 node-negative women with breast cancer for whom 10 or more years of follow-up was available, van’t Veer et al. [7] identified a gene expression signature, composed of 70 genes, that is predictive of rapid metastasis (‘poor prognosis’ signature). These results were validated in other retrospective studies [19-20]. According to these results, the 70-gene profiler predicted outcomes more accurately than all other clinicopathologic parameters. The functions of the 70 genes...
are associated with proliferation, invasion, metastasis, stromal integrity, and angiogenesis.

**Oncotype DX** • Also called 21-gene recurrence score assay. This is the first commercially available, RT-PCR-based test. It is performed in a Clinical Laboratory Improvement Act (CLIA) licensed centralized laboratory. It is currently designed as a prognostic and predictive assay to determine the 10-year risk for disease recurrence for women with ER-positive, lymph node-negative breast cancer. It is performed on formalin-fixed-paraffin-embedded samples. The test results report a recurrence score (RS) assigned into three risk categories: low, intermediate, and high. This test was developed by Paik et al. [21] who profiled archived breast cancer samples that were available from patients who participated in clinical trial NSABP B-20 and from patients from two other data sets. These investigators identified 21 genes, including five control genes; the differential expression of which was used to generate a recurrence score employing a constructed algorithm. The recurrence score was able to identify three groups of patients based on risk of recurrence with median follow-up of 10.9 years: low (recurrence rates of 10% or less), intermediate (recurrence rates of 10% to 30%), and high (recurrence rates of 30% or more). Similar results were obtained from a validation study on samples from a second NSABP trial, B-14, which has a median follow-up exceeding 14 years and includes node-negative, ER-positive patients [22]. These studies suggest that this assay predicts overall prognosis in addition to adjuvant therapy benefit in that group of patients and could identify a subset of patients, which would have excellent prognosis with hormonal therapy alone. The functions of the genes selected for this test are associated mainly with proliferation and estrogen receptors pathways.

**Breast Cancer Two-Gene Expression Ratio (H/I)** • This is an RT-PCR based assay testing the expression of the HOXB6 and IL17BR genes. It is performed on formalin-fixed-paraffin-embedded tissues. It predicts recurrence in patients with ER-positive, lymph node-negative primary breast cancer [23-26].

**Celera Metastasis Score** • This is an RT-PCR based assay testing the expression of 14 genes on formalin-fixed-paraffin-embedded tissues. It predicts recurrence in patients with ER-positive, lymph node-negative tumors. Preliminary studies indicate that this test predicts a 3.5-fold difference in risk for disease recurrence between the women at the highest risk and the women at the lowest risk [27].

**The Breast BioClassifier** • This is another RT-PCR based assay performed on formalin-fixed-paraffin-embedded tissues. It measures the expression of 50 genes to identify the different biological subtypes of breast cancer (luminal-A, luminal-B, HER2, and basal-like), provides a prognostic risk assessment and may identify groups of patients that may potentially benefit from personalized therapy [28].

**The Rotterdam Signature** • Also known as the 76-gene assay. This is a microarray-based performed on fresh-frozen tumor samples or tumors preserved into an RNA preservative solution. This assay is validated to predict prognosis in lymph node-negative patients independently of hormone receptor status [29-32].

**Invasiveness Gene Signature** • Another microarray-based test that measures the expression of 186 genes to predict prognosis in both node-negative and node-positive and both ER-negative and ER-positive patients [33].

**NuvoSelect** • This is a microarray-based assay that combines evaluation of several gene sets and claims to provide prognostic information as well as predictive therapeutic response to chemotherapy. It also provides ER and HER2 status [34-36].

**Cytochrome p450 CYP2D6 Genotyping** • ER-positive breast cancer patients with low or absent levels of CYP2D6 enzyme cannot activate tamoxifen and therefore would benefit more from aromatase inhibitors. Several assays using different technologies to determine CYP2D6 status are available, some are FDA approved [37-39].

The applications of gene expression profiling will only be expanding. Recently, a study reported that gene expression profiling could predict local recurrence after breast conserving surgery [40].

**GENE EXPRESSION PROFILING IN THE CLINICAL SETTING: LIMITING FACTORS AND PITFALLS**

Hundreds of markers have been tested and evaluated in thousands of publications, but only a few markers have been used successfully in clinical practice. Gene expression profiling has been the object of high profile studies and carries much hope; however, modern, complex and new techniques should undergo the same validation processes as the most basic ones. Adopting a test for routine clinical use requires standardization, proper validation, demonstration of diagnostic, prognostic, and predictive performance, demonstration of clinical benefit in routine clinical use and cost-effectiveness [41].

Although some tests are already commercially available, none of the gene expression profiling assays satisfies all these requirements and clinicians should be aware of the following limitations:

**Study design and proper validation**

Until now, tumor marker research has depended mostly on small studies with incomplete validations [42-43]. Current molecular profiling studies suffer from the same problems. It appears that the validation performance of several proposed signatures is inflated. Michiels et al. [44] showed that five of seven published molecular profiles do not actually perform better than chance. The first evaluation of the 70-gene signature showed near perfect accuracy, while the validation study revealed specificity of about 40% only [20].

According to Ein-Dor et al. [45], proper selection of appropriate genes in molecular profiling studies, requires sample size about 100-fold larger than the sample sizes
that have been used. Level I evidence for adopting new tumor markers requires analysis in prospective clinical trials designed to address the relevance of these markers in a clinical context [46-47]. None of the previous gene expression profiling studies is based on prospective trials. Up till now, only two of the current assays are being evaluated in prospective clinical trials and the first data will not be available before 2010 (see below).

Limitations of gene expression profiling data
Data from gene expression profiling are limited to specific tumor types and categories of patients. For example, Oncotype DX has been developed for estrogen-positive, lymph node-negative tumors only. Moreover, in the molecular classification, basal-like cancers have been associated with a less favorable outcome. But Weigelt et al. [48] showed that some special types of breast cancer associated with a good prognosis, such as medullary and adenoid cystic carcinomas, display also a basal-like profile. In addition, according to Esteva et al. [49] Oncotype DX failed to predict the risk of distant disease recurrence in patients with node-negative breast cancer who did not receive systemic therapy. Mina et al. [50] reported that Oncotype DX recurrence score did not correlate with pathologic complete response in patients, with newly diagnosed stage II or III breast cancer, who received preadjuvant chemotherapy.

Clinicians should be aware that data from gene profiling studies are limited to certain categories of patients and tumor types and cannot be applied in all clinical situations. Also, traditional clinicopathologic parameters should be integrated with data from gene expression profiling for proper patient’s management.

Standardization of gene expression profiling-based assays
Standardization of molecular tests is necessary to ensure that genes activity can be measured and analyzed in the same conditions in different patients. However, variability and therefore biases may derive from the sampling of the specimens. For example mRNA extracted from a breast cancer specimen with a healing biopsy site, from a previous intervention, can give an inaccurate high proliferative profile. The timing of processing after surgery, fixation, specimen preservation and the nature of the biopsy (core vs. excision) can impact gene expression profiling [51-53].

Cost
Gene expression profiling based tests are very expensive and many are available only in the companies that developed them.

CURRENT GUIDELINES AND PROSPECTIVE CLINICAL TRIALS

The 2007 ASCO guidelines [54] state that 21-gene recurrence score assay (Oncotype DX) can be used to predict the risk of recurrence in women with newly diagnosed, node-negative, ER-positive breast cancer who will be receiving tamoxifen and to identify patients who are predicted to obtain the most therapeutic benefit from adjuvant tamoxifen and therefore, not require chemotherapy. The guidelines also state that the present data are insufficient to recommend use of other assays (including the MammaPrint, the Rotterdam Signature and the H/I breast cancer gene expression ratio).

It appears however, that the true clinical relevance of these tests needs additional studies to be determined. Two ongoing clinical trials may bring some needed answers:

- **The TAILORx (Trial Assigning Individualized Options for Treatment) Trial**. This prospective clinical trial is sponsored by the National Cancer Institute and involves 900 sites in North America [55-56]. It will use Oncotype DX to guide treatment selection. This trial plans to enroll at least 10,000 women with estrogen or progesterone-positive HER2-negative, lymph node-negative breast cancer. Patients with low Recurrence Score will receive hormonal therapy alone, patients with high Recurrence Score will receive hormonal therapy and chemotherapy, and patients with intermediate Recurrence Score will be randomized into either hormonal therapy alone or hormonal therapy plus chemotherapy. Results from this trial will not appear before 2013.

- **The MINDACT (Microarray in Node-Negative Disease May Avoid Chemotherapy) Trial**. This prospective clinical trial is sponsored by the European Organization for Research and Treatment of Cancer and opened in August of 2007 [57]. It plans to compare the 70-gene Mammaprint assay against the standard clinicopathologic prognostic factors included in Adjuvant Online, in selecting node-negative breast cancer patients for adjuvant chemotherapy. Preliminary results from MINDACT may be presented in 2010.

CONCLUSION

Gene expression profiling in breast cancer has had important implications on our understanding of breast cancer. It will likely greatly impact prognostic assessment and prediction of response to therapy. However, despite the potential of this field and while the list of commercialized assays is expanding, several challenges remain and many issues need to be addressed. Will one assay prove to be superior to the others? Will newer and better panels emerge? Also, adoption of these assays by clinicians unfamiliar with their limitation is a serious concern and may cause harm to the patients if these tests are used in the wrong clinical setting. It is clear that large prospective studies and more clinical experience are needed to establish guidelines for gene expression profiling assays and additional studies are needed to confirm their scientific validity and their true clinical utility.

REFERENCES

2. Elston CW, Ellis IO. Pathological prognostic factors in

S. M. NASSER – Gene expression profiling in breast cancer

Lebanese Medical Journal 2009 • Volume 57 (2) 87