Who needs adjuvant systemic treatment? Predicting prognosis in node-negative breast cancer; from bench to bedside

Bueno de Mesquita, J.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Technology Insight:
tuning into the genetic orchestra using microarrays-limitations of DNA microarrays in clinical practice
Technology Insight: tuning into the genetic orchestra using microarrays—limitations of DNA microarrays in clinical practice

Ambreen Abdullah-Sayani, Jolien M Bueno-de-Mesquita and Marc J van de Vijver

Department of Pathology, Netherlands Cancer Institute, Amsterdam, the Netherlands


Summary

Scientific advances in the field of genetics and gene-expression profiling have revolutionized the concept of patient-tailored treatment. Analysis of differential gene-expression patterns across thousands of biological samples in a single experiment (as opposed to hundreds to thousands of experiments measuring the expression of one gene at a time), and extrapolation of these data to answer clinically pertinent questions such as those relating to tumor metastatic potential, can help define the best therapeutic regimens for particular patient subgroups. The use of microarrays provides a powerful technology, allowing in-depth analysis of gene-expression profiles. Currently, microarray technology is in a transition phase whereby scientific information is beginning to guide clinical practice decisions. Before microarrays qualify as a useful clinical tool, however, they must demonstrate reliability and reproducibility. The high-throughput nature of microarray experiments imposes numerous limitations, which apply to simple issues such as sample acquisition and data mining, to more controversial issues that relate to the methods of biostatistical analysis required to analyze the enormous quantities of data obtained. Methods for validating proposed gene-expression profiles and those for improving trial designs represent some of the recommendations that have been suggested. This review focuses on the limitations of microarray analysis that are continuously being recognized, and discusses how these limitations are being addressed.
Introduction

The discovery of 25,000 human genes as a result of the completion of the Human Genome Project has greatly accelerated the research into genotypic–phenotypic correlations in an aim to elucidate the functional taxonomy of genes in both normal tissues and disease states. The use of microarray technology to perform gene-expression profiling is an important adjunct to our current knowledge of genetics, because microarrays allow the study of thousands of genes simultaneously in a single, standardized, and cost-efficient experiment. Gene expression is a broad term used to describe the transcription of information encoded within DNA sequences into mRNA, and the subsequent translation of the mRNA information into proteins that regulate cell function. The gene expression of all cells is a highly dynamic process that alters in response to cell requirements and environment changes, and as a consequence of disease. A brief summary of the applications of gene-expression profiling by microarray analysis and additional application of microarray platforms are described in Table 1.

Table 1: Examples of current applications of microarray technology.

<table>
<thead>
<tr>
<th>Microarray type</th>
<th>Purpose</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression profiling</td>
<td>To determine the changes in gene expression that occur in response to different disease or normal states or to compare the gene-expression level in a cell over time</td>
<td>Tumor classification</td>
</tr>
<tr>
<td>SNP analysis</td>
<td>To detect mutations or polymorphisms in a gene sequence that are responsible for genetic variation</td>
<td>Development of prognostic and predictive markers, Drug development, Determination of genetic predisposition to a disease, Monitoring of disease progression, Discovery of DNA-based drug candidates, Tumor classification, Risk assessment, Development of prognostic and predictive markers, Evaluation of germinal mutations in individuals, Identification of somatic mutations in cancer</td>
</tr>
<tr>
<td>CGH</td>
<td>To identify genetic amplifications or deletions or to identify copy number changes in a particular gene</td>
<td></td>
</tr>
<tr>
<td>Resequencing arrays</td>
<td>To sequence portions of the genome</td>
<td></td>
</tr>
</tbody>
</table>

*Table adapted from the National Center for Biotechnology Information website. Abbreviations: CGH, comparative genomic hybridization; SNP, single nucleotide polymorphisms.

Gene-expression profiling has been used to study infectious and immunological diseases, but the predominant focus of microarray-based research has been in the study of cancer. Microarray analysis has made it possible to identify groups of genes according to their expression pattern or ‘genetic signature’, and this could potentially ameliorate the clinical management of a patient. Despite the vast amount of research in this field, a disparity exists between the quantity of publications and the quality of their interpretation. The quantity of articles that document the discovery of new gene profiles is as plentiful as the number of publications that scrutinize their interpretation. In this article we will briefly describe the methods of microarray-based gene-expression profiling and highlight the technique’s limitations, demonstrate the various levels at which both inter-experimental and intra-experimental variability can occur, summarize the suggestions that have been proposed to limit these problems, and describe the current status of genetic profiling in oncological practice. The different steps involved in gene-expression-profiling using microarray technology and the major limitations of this method are depicted in Figure 1.
Who needs adjuvant systematic treatment

Figure 1: Key considerations when performing gene-expression profiling using microarray analysis. The different steps in this process and their limitations are summarized. There are seven major steps in the development of a gene-expression signature using microarrays before it is ready for implementation as a test in clinical practice. The first step consists of defining the scientific question that should be answered by the output data. The accompanying hypothesis will help to construct a solid study design and to determine the type of samples and the applicable inclusion criteria (steps 2 and 3). Step 4 includes the actual microarray experiment by which the gene-expression data are obtained. The statistical analysis of the data generated in the experiment and the construction of the actual gene-expression profile is addressed in step 5. Step 6 includes the independent validation of the results, which is indispensable for the transition to step 7, the clinical implementation of the gene-expression profile. Figure modified with permission from reference 21 © (2002) Nature Publishing Group.

DNA microarray technology
A DNA microarray is an ordered arrangement of equidistant microscopic DNA spots attached to a solid surface, such as a glass, plastic or silicon chip. Hybridization is performed using corresponding probes that recognize and attach to the solid support; these can be complementary DNAs (cDNAs), oligonucleotides of varying length, or genomic sequences that are either radioactively or fluorescently labeled.
An array containing thousands of spots immobilized at predetermined locations can be generated by applying the DNA (e.g. cDNA or oligonucleotides) to the array using pins\(^1\) or inkjet technology,\(^4\) or by in situ photolithographic synthesis of oligonucleotides.\(^6\) For example, if inkjet technology is being used, an adapted color inkjet printer is used to synthesize oligonucleotides: each of the four color reservoirs is filled with a solution containing one of the four nucleotides from which DNA is made. The oligonucleotide sequences for each gene can then be assembled using a simple text file. Thus, the probes of a microarray are made of DNA and these probes are then used to detect the abundance of specific mRNA (transcripts) from the genes that correspond to the sequence of the probes on the array.

**Figure 2**: Gene-expression profiling using microarray analysis. The main steps involved and limitations of microarray analysis will be encountered. Two stages can be distinguished in microarray experiments: a pre-microarray experiment phase (tissue handling) and the microarray experiment phase (labeling of the RNA). In a microarray experiment the glass slide with the labeled DNA of interest plays a key role. As can be seen, a solid surface (in this example a glass microscope slide) contains thousands of spots. Each spot contains a large number of identical DNA fragments. Fluorescently labeled RNA from the samples are subsequently hybridized to the arrays. In this way, the amount of DNA fragments per spot indicates the expression level of a gene. The expression level of thousands of fluorescently labeled genes or spots on one microscope slide can be visualized with a fluorescent scanner. For each gene on the array, the amount of fluorescently labeled RNA bound represents the expression level of that gene in the tumor sample. The intensity of fluorescent signal can be measured and used in the statistical analysis. For each spot, the DNA fragments are derived from one specific gene. Figure courtesy of Dr R Kerkhoven. Abbreviations: DHFR, dihydrololate reductase; E2F1, E2F transcription factor 1; RB, retinoblastoma; SRC, sarcoma.
The concept of DNA or oligonucleotide arrays began in the mid-1980s. The first DNA arrays comprised nylon filters on a glass slide containing cDNA probes; these filters contained a much lower number of probes than the arrays that are currently used and were typically used with radioactively labeled targets. The introduction of pin-based robotic systems made it possible to dispense smaller volumes of DNA (approximately 150 microns) onto a glass slide, thus enabling a higher throughput system that represented one of the first microarrays.\(^6,7\) A summary of the developments in microarray-based gene-expression profiling is provided in Table 2.

### Table 2: The evolution of microarrays for gene-expression analysis over the years\(^5\)

<table>
<thead>
<tr>
<th>Type of array</th>
<th>Number of array spots</th>
<th>Target</th>
<th>Manufacturer</th>
<th>Year introduced</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nylon/oligonucleotide-cDNA</td>
<td>500-5,000</td>
<td>cDNA clones</td>
<td>NA</td>
<td>1996</td>
<td>Almost out of use</td>
</tr>
<tr>
<td>Glass-cDNA</td>
<td>50-40,000</td>
<td>cDNA clones</td>
<td>Affymetrix</td>
<td>1998</td>
<td>Rapidly decreasing</td>
</tr>
<tr>
<td>Glass/quartz—oligonucleotide</td>
<td>22,000-45,000</td>
<td>20mer oligonucleotides</td>
<td>Affymetrix</td>
<td>2000-2001</td>
<td>Increasing</td>
</tr>
<tr>
<td></td>
<td>22,000-44,000</td>
<td>60mer oligonucleotides</td>
<td>Agilent</td>
<td>2004</td>
<td>Increasing</td>
</tr>
<tr>
<td>Exon array</td>
<td>1 million exons</td>
<td>12mer oligonucleotides</td>
<td>Affymetrix</td>
<td>2005</td>
<td>Increasing</td>
</tr>
<tr>
<td>beadArray</td>
<td>50</td>
<td>50mer oligonucleotides</td>
<td>Illumina</td>
<td>2005</td>
<td>Increasing</td>
</tr>
</tbody>
</table>

\(^5\)Although nylon membranes provide some advantages, such as flexibility, low cost, and possibility of stripping and rehybridizations, their use has almost disappeared. Abbreviations: cDNA, complementary DNA; NA, not applicable.

Microarray analysis of cDNAs spotted onto glass slides was developed at Stanford University and has been used to study the expression levels of large numbers of mRNAs in cell lines and tumors.\(^8\) In brief, mRNA is isolated from cells and reverse transcribed in the presence of red-fluorescent-labeled nucleotides. The resulting fluorescent cDNA is then mixed with a green-fluorescent-labeled reference cDNA and the mixture is hybridized to the microarray. The reference mRNA is usually prepared from a mixture of cell lines, tumor samples, or normal tissues. Using a fluorescent scanner, the fluorescence level is digitized, and for each cDNA on the microarray the level of gene expression, relative to the reference, is determined and transferred to a linked computer database.

Similar methodology is used for oligonucleotide-based arrays. With this system of microarray analysis, the expression of each gene in a sample is also measured relative to the expression of the same gene in the reference mRNA. The photolithographic synthesis of oligonucleotides is a procedure developed by Affymetrix (Santa Clara, CA); for each gene, several different oligonucleotides are present on the array.\(^5\) In this technique, the hybridization on the array platform is carried out using RNA from the sample to be analyzed, without the use of a reference RNA; instead, oligonucleotides containing one mismatch in their sequence are used to correct for background hybridization. An array containing 25,000 probes can provide information on the expression of all genes present in the genome. For many genes, multiple splice variants have been identified, and for sufficient array detection an increasing numbers of probes are being developed.

Gene-expression profiles can provide an enormous amount of information on cell function; however, it is important to realize that these profiles can only be used to interpret cellular changes that affect mRNA synthesis. After translation of mRNA into
protein, many secondary protein modifications also play an important role in cell regulatory processes. Therefore, high-throughput analysis of proteins (‘proteomics’) will also contribute greatly to cancer research, an issue that was recently reviewed by Gulmann et al.9 Proteomic experiments can be tedious and have a lower throughput than RNA-based experiments;10 however, recent advances in proteomics have greatly increased the throughput of proteomic experiments.

Recently, small RNA molecules that do not code for proteins, such as microRNA, RNA interference, small interfering RNA (siRNA), and small modulatory RNA, have been discovered; these small RNAs have an important role in gene regulation.11 Genes that regulate the expression of molecules such as microRNA are found in the non-coding regions (or introns) of the genome. Although these RNAs do not code for proteins, they can control gene expression at the post-transcriptional level by degrading or repressing mRNA,11 thus influencing critical functions and biological processes.12 Although microRNA consists of only 1–5% of human genes (which equates to about 1,000 microRNA genes), each microRNA might regulate as many as 200 genes, which implies that over one-third of the genes that encode for proteins are regulated by microRNAs.11, 13 MicroRNA expression profiles can be used to classify human cancers12 because they reflect the developmental lineage and differentiation state of the tumors. Compared with normal tissue, microRNAs in tumors are generally downregulated and could thereby provide promising insight into tumor development and recurrence. For example, microRNA profiles are better predictors of metastatic origin for carcinomas of unknown primary (CUPs) than are conventional gene-expression signatures using mRNA.14

Statistical analysis of gene-expression data

Although not necessarily based on a gene-specific mechanistic hypothesis, good gene-expression profiling experiments should be planned and conducted with a clear objective, design and statistical analysis strategy.15 The main strategies used to identify categories of tumors by gene-expression profiling are unsupervised and supervised classification (Box 1). An important difference between these two methods is that for supervised classification clinical or pathologic information is used to find correlations with gene-expression patterns, whereas with unsupervised methods the tumors are grouped on the basis of their gene-expression pattern independent of the clinicopathological status.16, 17 A commonly used unsupervised method for examining gene-expression data is two-dimensional hierarchical cluster analysis. In this method, the tumors and genes are ordered according to similarities in their gene regulation, resulting in the clustering of tumors according to similarity in gene expression. This ordering often results in clustering of genes responsible for cellular processes, such as proliferation or inflammation. In this manner, several ‘signatures’ can be analyzed that reveal specific properties that the tumor cells and non-tumor cells contribute to the tumor mass. With time it is expected that an increasing number of signatures will be identified, making the correlation between gene-expression signatures in tumors and clinical behavior a powerful prognostic and predictive approach.

To find gene-expression patterns that can predict the clinical behavior of tumors (as in class comparison and class prediction), it is more appropriate to use a supervised
classification that can distinguish specimens on the basis of predefined clinical and pathologic information, because the classes are already known. Several methods for supervised classification have been developed, and although these techniques were not specifically designed for the analysis of gene-expression data they have already shown their usefulness for this purpose. Fundamental to each of these techniques is the identification of the pattern of expression of a combination of genes that can predict tumor behavior (e.g. the risk for development of distant metastases, or responsiveness to specific treatments).

A third classification method is functional annotation of gene-expression algorithms based on data obtained from in vitro and in vivo experiments. An example of such

Box 1 Definitions of terms used in microarray technology.

Unsupervised clustering (unsupervised classification)
The purpose of any clustering method is to group entities based on similarity of features. Gene-expression data derived from multiple microarray experiments, genes and tumor samples can be clustered on the basis of similar expression profiles. Hierarchical clustering produces a representation of the data in the shape of a binary tree, in which the most similar patterns are clustered in a hierarchy of nested subsets, resulting in the grouping together of tumors that are similar in their overall gene-expression profiles. Binary trees are two-dimensional diagrams also known as dendrograms, which illustrate the fusions or divisions made at each successive stage of analysis.

Supervised clustering (supervised classification)
Supervised clustering is a statistical method used to assess similarities between certain tumor types. This method can be defined as grouping of variables (e.g. genes) controlled by information about the Y variables (i.e. variable of interest), such as the tumor types of the tissues or clinical outcome of a disease. Supervised clustering is applied to classified samples with the objective of identifying clusters that have a high probability of sharing certain features characteristic of the class of interest.

Metagenes
Metagenes are linear combinations of individual gene-expression values, and have the potential to classify and predict cellular phenotypes resulting from deregulation of oncogenic pathways.

Adjuvant systemic therapy
A treatment that is used in addition to the primary therapy, such as surgery, to ensure that all microscopic disseminated cancer cells are destroyed. Adjuvant therapy for cancer can include cytotoxic chemotherapy and/or hormone therapy.
an algorithm is the 'wound response signature'. Chang et al. compared the gene-expression profile of fibroblasts exposed to 10% serum with fibroblasts growing at low serum conditions; this gene-expression signature was termed 'wound response signature' because of the fact that fibroblasts are only exposed to serum in a wound. Based on the expression pattern in this 'wound response signature', breast carcinomas can also be divided into 'wound response signature activated' and 'wound response signature quiescent' tumors. Other gene-expression signatures discovered using in vitro experiments can in a similar way be applied to gene-expression data in malignant tumors. Chang et al. have combined three algorithms—the molecular subtypes or portraits (derived from unsupervised hierarchical clustering of breast carcinomas), 'wound response signature', and the '70-gene signature'—to reveal links between wound healing and cancer progression in breast cancer. A gene-expression profile based on the wound response signature provides the basis for prospectively assigning a prognostic score to patients and can be scaled to suit different clinical purposes. The wound response signature improves risk stratification independently of known clinicopathological risk factors such as tumor size, histologic tumor grade and nodal status, and previously established prognostic signatures based on unsupervised hierarchical clustering (i.e. 'molecular subtypes') or supervised predictors of metastasis (i.e. the '70-gene signature').

Even when distinct subgroups are defined using gene-expression profiles, considerable heterogeneity within the broadly defined groups can confound the ability of the classifier to accurately predict outcomes for individual patients. One strategy by which to circumvent this heterogeneity is to combine or integrate multiple gene-expression patterns rather than a single gene-expression pattern, which provides a more powerful and robust classifier. Such defined sets of genes have also been termed 'metagenes' (Box 1); such metagenes can be combined with clinicopathological data to obtain optimal prognostic and predictive classifiers. Moreover, this framework provides a mechanism for combining multiple forms of data—both genomic and clinical—to characterize most individual patients' gene profile and achieve the goals of personalized treatment.

**Limitations of microarray technology**

The complexity of biological systems
Gene-expression profiling studies of human malignancies are highly complex, and the successful execution of such a study requires close collaboration between surgeons, pathologists, molecular biologists and bioinformaticists. The high-throughput nature of this technology combined with the expected plethora of data results in a high opportunity for errors. To ensure the accuracy and reliability of the resulting data, it is essential, therefore, that experiments are tightly regulated and quality controlled. At present, the sophistication of microarrays renders this a costly technology and consequently it is only available to specialized institutions. Microarray technologies, however, are rapidly improving and the costs of the technique continue to fall, thus paving the way for wider access and more-generalized usage.
The tissue sample
Frozen tumor tissue is indispensable for microarray experiments because the rapid fixation method ensures that the RNA quality remains optimal, with the result that subtle alterations in levels of differentially expressed genes can be accurately detected. Frozen tumor samples, however, are not widely available, because in most hospitals the tumor samples are directly fixed in formalin and embedded in paraffin blocks; storage of tissue in this way has been shown to result in RNA degradation. Paraffin-embedded tissue is not, therefore, ideally suitable for finding new gene-expression signatures, but may be more adequate for the validation of new gene-expression profiles for which the differentially expressed genes are already known. Experiments making use of mRNA have distinct problems. To begin with, there may be very small amounts of mRNA within the cell itself, which can be a significant problem in smaller sized tumors. mRNA is a very fragile molecule that can degrade within minutes of surgical manipulation, drastically affecting the interpretation of microarray data. Likewise, subtle variations in tissue handling and method of RNA extraction from samples can result in different levels of gene expression. Although there are no systematic studies of this subject, it is clear that measurements of specific mRNAs with a short half-life will be greatly affected by the interval between surgical removal and freezing of the specimen. These technical issues are also compounded by the added complexity of the heterogeneity of the tumor, and this heterogeneity expands as we consider the different individuals and populations that harbor the same genetically unstable tumor. In addition, the expression pattern in tumor cells is also determined by gene expression in the many other cell types present in varying degrees in clinical tumors, such as fibroblasts, endothelial cells, and inflammatory cells. Rigid adherence to experimental technique, consistency in the timing and sampling of tissue, careful documentation of the amount of tumor cells in a sample and duplication of the experiment using a single reference RNA are all means by which sources of error can be eliminated. Technological advances also enable standardization of this process.

The search for differentially expressed genes
The hunt for differentially expressed genes involves a number of steps, including chip production, probe hybridization, image quantification, normalization and data interpretation, before the biological question can be addressed. Until a consensus is reached to standardize each process, inter-experimental variability will remain commonplace, inhibiting the transfer of microarrays from the bench to the bedside. Pitfalls that clinicians should be aware of include unstable molecular signatures and gene misclassification. The list of genes included in a molecular signature (based on a training set and the proportion of misclassification seen in one validation set) depends not only on the statistical methods (discussed in the next section), but also to an even larger extent on the selection of the patients in the training sets. Michiels et al. found that five of the seven largest published studies addressing cancer prognosis did not classify patients better than chance. Another important issue is that results from different microarrays are expressed as levels in relation to a nonstandardized reference RNA. It has been recognized that the implementation of standardized units for the level of
gene expression will be an important advance towards regulation of microarray data.\textsuperscript{2} Not all microarray problems, however, are of such a large scale. Simple experimental errors such as the mishandling of plates can give rise to subsequent errors in data. The importance of tracking all levels of an experiment to allow evaluation and correction (if needed) has been shown by investigators in Stanford University who created an algorithm called MuFu (MixupFixup) to facilitate recognition of production errors prior to hybridization.\textsuperscript{26} Technical errors in an experiment can give rise to missing data, and this will eventually affect the statistical outcome of a study.\textsuperscript{35} Tools such as LinCmb,\textsuperscript{35} GMC (Gaussian mixture clustering),\textsuperscript{36} and BPCA (Bayesian principal component analysis)\textsuperscript{37} have been developed to input the missing values of an experiment and improve the significance of a study.

Statistical analysis

The proposed association of a genetic signature with disease outcome has been observed to be significantly stronger in preliminary studies than in subsequent research.\textsuperscript{38} This outcome can be partially explained by ‘overfitting’, which is one of the major limitations of supervised clustering methods. Overfitting indicates that the number of parameters in a model is too great relative to the number of cases or specimens available. Because the gene-expression pattern is optimized to predict tumor behavior, the model will fit the original data but might predict poorly for independent data. Consequently, it is essential to obtain an unbiased estimate of the true error rate of the predictive power or accuracy of a gene-expression pattern. Methods for obtaining improved predictors include ‘leave-one-out’ cross-validation; with this method, a gene-expression predictor for a specific clinical behavior is built by leaving out one or more samples. Cross-validation involves repeatedly splitting the data into a training set containing most of the samples and a test set containing the remaining samples, and this process is repeated for the many training-set and test-set partitions. The resulting error rates observed for the test sets are then averaged, and the predictions for each test set are unbiased estimates of the true prediction error because the test-set samples are not used in the development of the model used for their prediction.\textsuperscript{39} Subsequently, it is tested whether the gene-expression pattern correctly predicts the clinical behavior of the left-out sample(s). A predictor that is produced as a result of a small, properly cross-validated error rate for a collection of tumor specimens is a potentially important finding, but one that still requires further validation.\textsuperscript{39} Although in studies with a small sample size (which can be arbitrarily defined as <50 cases) cross-validated error estimates are nearly unbiased, they tend to have large variance and confidence intervals. A study with a sample size in the thousands is better equipped to understand the degree of biological variation within an experiment,\textsuperscript{35} and the resulting genetic signature will also have greater statistical significance and clinical applicability.\textsuperscript{40} It is essential to validate a predictive gene-expression pattern in a sufficiently large independent series of patients. Unfortunately, a major limitation is the unavailability of large amounts of frozen tumor samples. One limitation of unsupervised cluster analyses is that it provides qualitative and not statistically valid quantitative information on differences in expression level between genes or classes. Although unsupervised clustering gives insight into the differentially expressed genes (quality),
it does not quantify the level of upregulated or downregulated expression (quantity). Many authors make the mistake of performing Kaplan–Meier statistical analyses on series of samples that do not belong to a cohort of unselected patients. For a good and profound validation of prognostic or predictive gene-expression profiles, an unselected patient cohort is indispensable.

The applicability of microarrays in clinical practice

Microarrays have been used to study several tumor types, most notably breast, ovary, colon, gastric, leukemias, malignant lymphoma, prostate, lung, CUP, and malignant melanoma. Analysis and accurate clinical interpretation from the currently available data is a challenging task, as there are numerous inter-experimental variations that can significantly influence the interpretation of results. Proper statistical analysis and independent validation in large series of patients is the only way to address this dilemma. To understand the current status and relevance of the gene-expression profiles that have been developed, we have chosen to highlight a few important examples. Since current clinical trials evaluating the use of microarrays in the management of patients with breast cancer are ongoing, we have chosen to separate this subject for purposes of detailed discussion in a separate section.

Hematologic malignancies

Acute leukemia is divided into histologic subgroups of clinical relevance—for example, lineage (B-cell or T-cell) and molecular subtypes—that differ with respect to clinical outcome and response to treatment. Current methods of cytogenetic and morphological analysis, however, are tedious and not widely available. Microarrays can revolutionize this diagnostic process, because all relevant histopathological abnormalities can be assayed in a more standardized manner. Recent studies using microarrays have revealed more-refined classification of disease subtypes than provided by currently used methods.

Yeoh et al. studied 360 cases of pediatric acute lymphoblastic leukemia. Using unsupervised hierarchical cluster analysis, these researchers were able to correctly identify the leukemia subtypes of prognostic significance, namely T-cell acute lymphoblastic leukemia, E2A-PBX1, BCR-ABL, TEL-AML, mixed-lineage leukemia rearrangement, and hyperdiploid with more than 50 chromosomes. The authors also identified another subgroup with a unique genetic signature, the significance of which is yet unknown. The same group was able to develop a class discriminator that had a diagnostic accuracy of 97% upon reanalysis of their data. This gene-expression profile holds great promise in clinical practice, and efforts should be made to refine the signature for general application. Gene-expression profiling is not only aimed at better classifying disease entities, but can also lead to the identification of novel targets for therapy. For example, mixed-lineage leukemia cells have distinctly elevated levels of the receptor tyrosine kinase FLT3, and thereby represent an opportunity for targeted therapy. Armstrong et al. studied the effect of the small molecule inhibitor of FLT3, and showed that it halted tumor progression.
Gene-expression profiling of diffuse large B-cell lymphoma (DLBCL) has allowed the categorization of this cancer into groups based on cellular origin; in one group the gene expression was characteristic of germinal-center B cells (i.e. germinal-center B-like DLBCL), whereas the other group consisted of genes normally induced during in vitro activation of peripheral blood B cells (i.e. activated B-like DLBCL). Patients with germinal-center B-like DLBCL have significantly better overall survival (5-year overall survival of 76%) than those with activated B-like DLBCL (5-year overall survival of 16%; P <0.01); this knowledge can allow stratification of patients for clinical management. Attempts at validation of these results have shown that the BCL6 and HGAL (GCET2) genes are specifically expressed in the germinal-center B cells and predict overall survival, while other differentially expressed genes, such as CD10 (also known as MME), do not predict for overall survival.

In 2004, Lossos et al. used information from microarray studies to assess a set of six genes that were predictive of outcome in patients with DLBCL using reverse transcription polymerase chain reaction (RT-PCR) assays. This is the same group that previously defined the active and germinal-center B-cell prognostic groups; however, in this current study these researchers showed that only two genes from the germinal-center B-cell signature—LMO2 and BCL6—and three genes from the activated B-cell signature—BCL2, CCND2 and CCL3 (previously named SCYA3)—had any predictive value.

Prostate cancer
Prostate cancer is an extremely difficult disease to judge clinically. The prevalence of this cancer is vast, and in men over 80 years of age a prevalence of as high as 80% has been reported in autopsy series. The majority of these men will die from causes unrelated to the cancer itself, indicating that many of these tumors are indolent. The introduction of prostate-specific-antigen screening has increased the identification and therefore the perceived incidence of early-stage prostate tumors, many of which are destined to remain indolent. Identifying patients with indolent tumors can spare many unnecessary radical treatment and its associated morbidity. Yu et al. studied 152 prostate tissue specimens, including frozen tissue samples from tumor, adjacent non-tumor and donor tissue. Using an Affymetrix® platform (Affymetrix, Inc., Santa Clara, CA), they analyzed 37,777 probe elements, and, using a combination of principal component analysis, supervised hierarchical clustering and 10-fold cross-validation, they developed a 70-gene expression profile predictor of aggressiveness with an accuracy of 93%. The results were independently validated in a small group of 23 patients.

In the same year, Lapointe et al. studied 121 frozen tissue samples that comprised 62 cases of prostate cancer, 41 normal tissue specimens, and 9 specimens of lymph node metastasis. Using a GenePix® platform (Astros Acquisition Sub II, LLC, Sunnyvale, CA), they studied 39,711 gene probes and used unsupervised hierarchical clustering to analyze the groups. These investigators were able to distinguish malignant from normal tissue, and further categorize the malignant tumors into groups based on risk of recurrence. They were able to streamline immunohistochemical analysis for MUC1 on 225 independent prostate tumors, and showed that elevated levels of MUC1 were correlated with aggressiveness of the prostate tumors in the training set.
and concurrent high risk of recurrence in the independent validation set ($P= 0.003$). These data show that gene-expression profiling cannot only identify those genes that discriminate indolent from aggressive tumors, but can also quickly screen for genes of interest, which can be adapted to widely available clinical tests such as immunohistochemistry. Although gene-expression profiling in prostate cancer is possible, it is not yet ready for clinical use as its validation is still lacking.

Carcinoma of unknown primary
The accurate diagnosis and targeted management of patients with (adeno)CUP (also called tumors of unknown origin) can sometimes influence patient survival. Standard histopathological measures are often incapable of solving the problem of diagnosing tumors of unknown origin, and this makes CUP a very attractive area for gene-expression profiling, as this technique could reveal more information about the tumor origin and help define suitable treatment options. At present, no literature on gene-expression profiling using microarrays in CUPs is available, although recent studies using Serial Analysis of Gene Expression (SAGE) have been able to generate expression patterns that can be used to trace the origin of the tumors. Buckhaults et al. used SAGE to select a group of genes that could substantially discriminate between tissues of different origins. A five-gene panel was developed that could distinguish between ovarian, breast and colon tumors and pancreatic adenocarcinoma, with a predictive accuracy of 81%; this model was validated in an independent series of 62 samples using quantitative real-time PCR. One of the major problems with SAGE analysis is that huge effort is required to study just one sample, thereby limiting the applicability of the technique for clinical use. Companies such as Arcturus and Agendia have already established commercially available gene-expression profiles using paraffin-embedded material, but rigorous scientific experimentation and validation is required for proper placement of this test in clinical practice. It is likely that microarray tests based on microRNAs will be superior to those based on mRNA.

Gastrointestinal tumors
The ability to accurately assess a patient’s risk of metastasis can profoundly alter clinical management. This statement is especially true for intermediate-stage tumors such as colon cancers of Dukes stages B for which treatment options are not clear. A patient categorized into a high-risk group can undergo treatment with adjuvant systemic therapy (Box 1); patients identified as low risk can be spared the side effects of an unwarranted treatment. Eschrich et al. studied 78 colon cancer specimens, and developed a 43-gene prognostic classifier that could predict with 90% accuracy the likelihood of survival at 36 months (93% sensitivity; 84% specificity). These results hold promise for tailored therapy for individual patients, which is currently not possible on a biological level. Unfortunately, we feel that not enough work has been done in this important field as very few relevant articles can be found in the literature, and thus more promising research in this clinically relevant field is warranted. For patients with esophageal carcinoma, lymph-node metastasis is a determinant of the therapeutic strategy. Current methods such as CT scanning and endoscopic ultrasound, however, do not provide accurate information on patient status.
Kan et al.\textsuperscript{60} studied 28 primary esophageal squamous cell carcinomas and applied a supervised classification technique called ‘artificial neural networks’ to develop a gene signature that showed 86\% accuracy in predicting lymph-node metastasis. Tamoto et al.\textsuperscript{61} performed a similar analysis using 36 esophageal tumor specimens. These investigators developed a 44-gene signature that was predictive of lymph-node metastasis in approximately 90\% of cases. Some genes present in this signature are known to have a biological role in metastasis. Like many current microarray experiments, the differences occur in similar types of studies because of a lack of consensus amongst the researchers. For example, in the above-mentioned studies in esophageal cancer, one pertinent difference is that Kan et al.\textsuperscript{60} used normal esophageal tissue as a reference RNA, whilst Tamoto et al.\textsuperscript{61} used tumor tissue. An apparent problem with current microarray studies is that the results have been obtained from relatively small groups of patients. Studies can have discordant data because the sample size in either the study or reference group is too small to allow the generation of a classification scheme that represents the biological complexity of a tumor.

Breast cancer
Results form gene-expression profiling studies for breast cancer are gradually being implemented in the clinic. The fact that prognostic factors have been used for the last 20 years to guide adjuvant systemic treatment of breast cancer means that gene-expression profiling can become an important adjunct to the known prognostic factors. For breast cancer, three relevant gene-expression profiles associated with prognosis have been identified: a 70-gene classifier,\textsuperscript{21, 22} a 21-gene signature,\textsuperscript{41} and a 76-gene expression profile.\textsuperscript{42} Table 3 shows the relevant similarities and differences between the three signatures. New additional signatures are being developed all the time, such as the recently identified signature from Pawitan and colleagues.\textsuperscript{62}

In 2002, a 70-gene prognostic signature was identified by the Netherlands Cancer Institute in Amsterdam.\textsuperscript{21, 22} The investigators initially studied the frozen samples of 78 tumors on an Agilent\textsuperscript{®} platform (Agilent Technologies, Inc., Palo Alto, CA; 25,000 probes).\textsuperscript{21} Using supervised classification, they found that in patients under the age of 53 years the expression of a set of 70 genes best correlated to distant metastases as a first event in node-negative cancer. This 70-gene expression signature was subsequently validated in a second partially independent validation series of 295 breast cancer patients from the same institute.\textsuperscript{22} For the node-negative patients in the validation series, 61 tumor samples overlapped with the training series. The 70-gene signature also proved to correlate with the outcome of 144 node-positive breast cancer patients (10-year overall survival rates for ‘poor’ and ‘good’ prognosis signature were 59.5\% and 92.0\%, respectively).\textsuperscript{22} This 70-gene signature outperformed the traditional clinicopathological features, with a hazard ratio of 4.6 (95\% CI 2.3–9.3, \(P < 0.001\)) in a multivariable analysis. A ‘good’ prognosis signature was present in 39\% of the patients, and was associated with a 94.7\% probability of freedom from distant metastases in the first 10 years, and an overall survival of 97.4\% independent of nodal status. A ‘poor’ signature was present in 61\% of
patients, and at 10 years 60.5% of these patients remained free of distant metastasis and the overall survival was 74.1%. A gene-expression-profiling-based commercial test, called MammaPrint®, is offered by Agendia (Amsterdam, The Netherlands). In 2004, the American National Surgical Adjuvant Breast and Bowel Project (NSABP) in cooperation with the company Genomic Health identified a recurrence score including 21 genes that quantified the likelihood of distant recurrence in tamoxifen-treated patients with node-negative, estrogen-positive breast cancer.41 Fixed, paraffin-embedded tumor tissue was assessed and gene-expression measured,63 resulting in the development of the Oncotype DX® assay (Genomic Health, Inc., Redwood City, CA). The list of 21 genes (16 cancer-related genes and five controls) in this assay and the recurrence-score algorithm were designed by analyzing data from three independent preliminary studies involving 447 patients and 250 candidate genes identified in earlier (including microarray-based) studies.

Table 3: Comparison between the three classified expression profiles: the 70-gene classifier,21, 22 the 76 gene expression profile,42 and the 21-gene signature41.

<table>
<thead>
<tr>
<th>Article</th>
<th>A gene-expression signature as a predictor of survival in breast cancer</th>
<th>Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer</th>
<th>A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>van de Vijver et al.21 and van’t Veer et al.21</td>
<td>Wang et al.42</td>
<td>Park et al.43</td>
</tr>
<tr>
<td>Publication year</td>
<td>2002</td>
<td>2009</td>
<td>2004</td>
</tr>
<tr>
<td>Medical centre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial partner</td>
<td>MammaPrint®</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature of specimens used for microarray</td>
<td>Frozen tissue</td>
<td>Frozen tissue</td>
<td>Paraffin-embedded</td>
</tr>
<tr>
<td>Patient cohort (validation)</td>
<td>10,563</td>
<td>22,215</td>
<td>666</td>
</tr>
<tr>
<td>Patient age</td>
<td>&lt;53 years</td>
<td>50–70 years</td>
<td>All</td>
</tr>
<tr>
<td>Tumor size</td>
<td>50 mm</td>
<td>10–100 mm</td>
<td>All</td>
</tr>
<tr>
<td>Node status</td>
<td>Node-negative/node-positive</td>
<td>Node-negative</td>
<td>Node-negative</td>
</tr>
<tr>
<td>ER-positive patients</td>
<td>77%</td>
<td>73%</td>
<td>100%</td>
</tr>
<tr>
<td>Patients treated with AST</td>
<td>Chemotherapy 7%, hormonal therapy 7%, both 7%</td>
<td>No adjuvant systemic treatment 100%</td>
<td>Breast-cancer treated (no chemotherapy)</td>
</tr>
<tr>
<td>Median follow-up time</td>
<td>6.4 years (range 3.0–15.3 years)</td>
<td>6.4 years (range 5.0–14.5 years)</td>
<td>10 years</td>
</tr>
<tr>
<td>Patients with distant metastasis</td>
<td>36%</td>
<td>25%</td>
<td>15%</td>
</tr>
<tr>
<td>Number of genes in signature</td>
<td>70</td>
<td>76</td>
<td>21</td>
</tr>
<tr>
<td>Primary outcome measures</td>
<td>Distant metastasis-free survival</td>
<td>Distant metastasis-free survival</td>
<td>Disease-free survival</td>
</tr>
<tr>
<td>Patients with a low-risk distant-metastasis-free survival of patients (LMFS)</td>
<td>At 10 years: 95.6%</td>
<td>At 5 years: 96.0%</td>
<td>At 10 years: 96.0%</td>
</tr>
<tr>
<td>OS of poor or high risk profile (OS)</td>
<td>At 10 years: 82.2%</td>
<td>At 5 years: 80.0%</td>
<td>At 10 years: 64.5%</td>
</tr>
<tr>
<td>OS of good or low risk profile (OS)</td>
<td>At 10 years: 84.7%</td>
<td>At 5 years: 80.0%</td>
<td>At 10 years: 64.5%</td>
</tr>
<tr>
<td>HR multivariable analysis (OR; CI; P value)</td>
<td>4.9 (2.3–9.3; P &lt; 0.001) (poor vs. good)</td>
<td>Cell death, cell cycle, proliferation, DNA replication and repair, immune response</td>
<td>2.61 (1.53–4.45; P &lt; 0.001) (high vs. low)</td>
</tr>
</tbody>
</table>

*If patients in the validation series20 were also part of the training series.21 Abbreviations: AST, adjuvant systemic treatment (chemotherapy and/or hormonal treatment); EIC-DOH, Erasmus Medical Center-Corabel van den Hout; Vioxx: ER, estrogen receptor; HR, hazard ratio; MDA, Merox Sharp & Domccie; NA, not applicable; NSABP, National Surgical Adjuvant Breast and Bowel Project; OS, overall survival.
The 16 cancer-related genes were selected based primarily on the correlation of their performance with outcome in three trials—one of the trials is the NSABP B-20 trial, but the other two are not specified. Detailed descriptions of how the 16 genes were selected are lacking; the weight attributed to each of the 16 genes differs, but how the investigators arrived at these different weights has not been described in any detail.

To test the prognostic value of the recurrence score, RT-PCR was successfully used in 668 paraffin-embedded tumor blocks out of a larger study population of tamoxifen-treated patients in the NSABP B-14 study. Tumor and patient characteristics in this subpopulation were similar to those of the total study population. The patients were of all ages and had tumors that were pT1–T2 estrogen receptor (ER)-positive, and were treated with tamoxifen. Using this 21-gene recurrence score, 338 patients (51%) had a low-risk, 22% an intermediate-risk, and 27% a high-risk profile. Multivariable analysis showed that the hazard ratio of this recurrence score was 2.81 (95% CI 1.70–4.64, \( P < 0.001 \)). Clinicians in the US are using the 21-gene recurrence score in clinical practice based on validation studies described on the Genomic Health website (www.genomichealth.com). Only one study, however, has been published in the literature, and ironically this study failed to validate the recurrence-score algorithm.

In 2005, the Erasmus Medical Centre in Rotterdam in cooperation with the American company Veridex identified a signature of 76 genes that could identify node-negative breast cancer patients at high risk of distant recurrence who were therefore eligible for adjuvant systemic therapy. Using a retrospective study design, the Affymetrix chip U133A, which contains 22,000 genes, was used on frozen samples of 286 untreated node-negative T1–T3/4 breast cancer patients of all ages. This prognostic signature was identified using a training series of 171 tumors and consists of two separate profiles, one for ER-positive (60 genes) and one for ER-negative (16 genes) breast carcinomas. The gene-expression levels were analyzed using the log rank test, and validated in an independent set of 115 tumors without any overlap with the training set. After 60 months, the distant-metastasis-free survival in the 65% of patients with a ‘poor’ profile was 53%, and this number dropped to 49% after 80 months. A ‘good’ profile was observed in 35% of patients, the disease-free survival of whom was 93% after 60 months, and 88% after 80 months. After 60 months the patients with a ‘poor’ profile had an overall survival of 70%; the rate decreased to 63% after 80 months. Conversely, for patients with a ‘good’ profile the overall survival rates for 60 months and 80 months were 97% and 95%, respectively.

Although the 70-gene profile from the Amsterdam group and the 76-gene profile of the Rotterdam group have only three genes in common, most genes are involved in the same regulatory pathways. A reason why the overlap between these two signatures is small is that different microarray platforms were used (i.e. Agilent® and Affymetrix®). Among the differences between these two platforms is the fact that different DNA probes are used for the detection of specific mRNAs—the Agilent® platform is a dual color hybridization using a reference RNA for each sample, whereas the Affymetrix® platform only uses RNA from the tumor sample.

There are disadvantages in assessing gene-expression profiles identified from
frozen tumor tissue. First, frozen tumor tissue is not widely available in clinical practice, and second, the assay is difficult to perform. These factors give the RT-PCR assay of Genomic Health an advantage over the other assays, since this assay can be performed more easily on readily available paraffin-embedded material. For samples that can be analyzed using paraffin-embedded tumor tissue, validation of the prognostic value can be established retrospectively from clinical trial material. For tests requiring frozen material, prospective studies are needed for validation. At present, the 70-gene signature\textsuperscript{21,22} is being used in The Netherlands in the RASTER trial (Netherlands Cancer Institute collaboration with the Dutch Health Care Insurance Board), and soon the MINDACT trial (TRANSBIG consortium in collaboration with the European Organisation for Research and Treatment of Cancer) is expected to start using this signature.

All three signatures outperform classical clinicopathological parameters in a multivariable analysis, but insufficient independent validation has been performed to assess the reproducibility of the results.\textsuperscript{21,22,41,42} It will be interesting to determine whether or not the three profiles categorize patients into the same prognostic groups. Although gene-expression profiling using microarray technology is very promising, there are still issues regarding its clinical applicability, and it is not yet known whether these profiles will genuinely improve the selection of those patients best eligible for adjuvant systemic treatment.

A key concern is when these gene-expression-based prognostic classifiers will be ready for clinical use. In the past 15 years there have been hundreds, even thousands, of studies attempting to identify novel prognostic factors in breast cancer, often employing molecular techniques. The factors used to predict prognosis and guide adjuvant treatment in most clinical guidelines are nodal status, tumor size, histologic grade, patient age and hormone-receptor status; recently, HER2 status has also been included. The reason that none of the factors has been implemented in clinical practice is that validation of the prognostic value often failed, and few prospective studies were properly planned. It appears that the gene-expression-based classifiers are more appealing to clinicians and patients, leading to a pressure to implement them in the clinic as soon as possible. Prospective validation of each of these classifiers in sufficiently large representative patient cohorts, however, is required before these tests can be used in the clinic. Premature use of these tests could lead to inappropriate advice on adjuvant systemic treatment, resulting in both undertreatment and overtreatment of patients with breast cancer.

**Future directions**

Microarray analysis is a rapidly evolving technology that has yielded enormous quantities of data, which are not yet fully understood. Leading scientific journals require investigators carrying out DNA microarray research to deposit their data in an appropriate international database,\textsuperscript{65} following a set of guidelines (MIAME guidelines, or Minimum Information About a Microarray Experiment\textsuperscript{2} Table 4). This approach offers an opportunity to propose alternative analyses for these data. Michiels et al. took advantage of this repository to analyze different datasets from published studies of gene expression as predictors of cancer outcome.\textsuperscript{34} They observed that
most molecular signatures and misclassifications are unstable, and their main conclusion was that the list of genes included in a gene-expression profile (based on one training set and the proportion of misclassifications seen in one validation set) depends greatly on the selection of patients in the training sets.

Table 4: Useful websites with URLs for microarray-based research.

<table>
<thead>
<tr>
<th>Website Type</th>
<th>Website/URL</th>
</tr>
</thead>
</table>
| Information        | Microarray Gene Expression Data (MGED) Society  
|                    | The Jackson Laboratory, Gary Churchill's Statistical Genetics Group        | [www.jax.org/staff/churchill/leabite](http://www.jax.org/staff/churchill/leabite) |
| Protocols          | Broad Institute Cancer Program                                             | [www.broad.mit.edu/cancer](http://www.broad.mit.edu/cancer) |
|                    | TMA Microarray Software Suite                                              | [www.tm4.org](http://www.tm4.org) |
|                    | ArrayExpress database (EBI)                                                | [www.ebi.ac.uk/arrayexpress](http://www.ebi.ac.uk/arrayexpress) |
|                    | Cancer Profiling Database Oncomine                                          | [www.oncomine.org/main/index.jsp](http://www.oncomine.org/main/index.jsp) |
|                    | GenesX-Lite (NCGR)                                                         | [www.ncgr.org/genex](http://www.ncgr.org/genex) |
|                    | HuGE Index.org                                                             | [www.biotechnologycenter.org/hio](http://www.biotechnologycenter.org/hio) |
|                    | Microarray Gene Expression Data (MGED) Society                             | [www.mged.org](http://www.mged.org) |
| Commercial         | Affymetrix                                                                 | [www.affymetrix.com](http://www.affymetrix.com) |
|                    | Agilent                                                                    | [www.agilent.com](http://www.agilent.com) |
|                    | Artusis                                                                    | [www.arctur.com](http://www.arctur.com) |
|                    | Agendia                                                                    | [www.agendia.com](http://www.agendia.com) |

With the recognition of the limitations of microarray analysis, but also of the measures that can be taken to reduce these limitations, and an understanding of the power of this technology, there is no reason why microarrays should not become a useful, if not essential, clinical tool. Until optimal gene-expression profiling is suitable for the clinic, the prognostic value of published microarray results in cancer studies should be interpreted with caution.

Technical issues should be addressed before this technology is translated into clinical practice. New methods of data analysis will increase the possibility of finding reliable prognostic and predictive profiles. Reproducibility and quality control are essential issues when using this technology as a standard assay in hospitals. In our opinion, the best way to improve analysis of microarray data is to start out with a clear study objective (e.g. class comparison, class prediction or class discovery study). If a prognostic or predictive profile is being built (class comparison or prediction), studies should be performed with statistical rigor and be reported clearly and with unbiased statistics. Usage of the REMARK tumor guidelines for reporting of tumor marker studies could be a good start. Currently, many studies have included only a small
number of patients; however, for the future, adequate study size in combination with appropriate clinical design, adjustment for known predictors and proper validation will yield more reliable results and less overestimation of the predictive value of a classifier. These features are essential in order for this highly promising technology to move from bench to bedside.

Simon et al. recommend that supervised methods rather than cluster analyses be used for class prediction and class comparison studies, as cluster analyses are less powerful than supervised methods for distinguishing predefined classes, and do not provide valid statistical identification of differentially expressed genes. If cross-validation is being used to estimate the prediction accuracy, the entire model-building process that includes selection of informative genes should be repeated for each cross-validation training set. Even after a successful cross-validation strategy is used, independent validation in an unselected patient cohort is indispensable. If a separate dataset is being used for validation, it should be an unselected cohort that is sufficiently large to provide meaningful confidence intervals for prediction accuracy. Inadequate validation leads to the publication of overoptimistic results. Finally, the investigators should be urged not to make strong claims about the value of new prediction algorithms without comparing them against standard prediction methods.

Conclusion

Microarray technology is a very promising and quickly developing area of expertise that is leading to the discovery of numerous new signatures in a very short period of time. Although the use of gene-expression profiles in clinical practice is very appealing, we should be very cautious in interpreting all the high-throughput data and results before we start using any of the identified gene-expression signatures in daily clinical practice. Lack of frozen material from patients that are treated has been an important limiting factor preventing proper validation of reported gene-expression profiles. An ideal way to solve this problem is by incorporating gene-expression profiling into prospective randomized clinical trials. With proper validation and standardization of the microarray process, we expect that tests based on discoveries through gene-expression profiling will gradually enter the clinic in the next 10 years as an important adjunct to current clinical parameters.
Key points
- Gene-expression microarrays are tools that can be used to simultaneously measure the level of gene expression of all genes within a cell
- Recent advances in technology and science are propelling microarray-based tests into the forefront of medicine
- In the last 5 years, numerous gene-expression profiles predicting prognosis and response to specific therapies for several malignancies have been reported
- Current limitations in the technology as well as in the biostatistical methods of analysis of the large quantities of data generated from microarray tests predispose the results to misinterpretation
- Standardization of protocols and validation of current profiles will have to ensure that gene-expression profiles are reliable and reproducible, and therefore ready for clinical implementation

Competing interests
The authors declared they have no competing interests.

Acknowledgments
We would like to thank L Wessel for advice and critical reading of the manuscript. We thank R Kerkhoven for his artwork for Figure 2.

Review criteria
The information for this Review was compiled by searching the PubMed and MEDLINE databases for articles published up to 15 November 2005. Electronic early-release publications were also included. Only articles published in English were considered. The search terms used included “DNA microarray” in association with other search terms: “reviews”, “RNA”, “DNA”, “gene expression profiling”, “classifier”, “molecular markers”, “tumor markers”, “cancer biomarker”, “prognosis”, “prognostic profile”, “predictive profile”, “multivariate analysis”, “classification”, and “risk assessment”. References were selected based on the best clinical or laboratory evidence, especially if the work had been corroborated by published work from other centers. Priority was given to studies in high impact factor journals when available. When possible, primary sources have been quoted.
Who needs adjuvant systematic treatment

References

Who needs adjuvant systematic treatment