VOLUME 23 · NUMBER 8 · MARCH 10 2005

JOURNAL OF CLINICAL ONCOLOGY

REVIEW ARTICLE

Gene Expression Profiling of Breast Cancer: A New Tumor Marker

Laura J. van't Veer, Soonmyung Paik, and Daniel F. Hayes

J Clin Oncol 23:1631-1635. © 2005 by American Society of Clinical Oncology

From the Division of Diagnostic Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands; Division of Pathology, National Surgical Adjuvant Breast and Bowel Project, Pittsburgh, PA; and the Breast Oncology Program, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI.

Submitted November 14, 2004; accepted December 7, 2004.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Daniel F. Hayes, MD, Breast Oncology Program, University of Michigan Health System, 1500 E. Medical Center Dr, CCGC 6203, Ann Arbor, MI 48109-0942; e-mail: hayesdf@umich.edu.

0732-183X/05/2308-1631/$20.00

DOI: 10.1200/JCO.2005.12.005

INTRODUCTION

In ideal clinical oncology practice, one would like to have a single platform assay that can provide both prognostic (estimates of risk for failure after surgery alone) and predictive (estimate of benefit from specific therapy) information. Despite years of research, only estrogen receptors (ER), progesterone receptors, and HER-2 have been widely accepted for routine use in breast cancer, serving as predictive factors for endocrine and trastuzumab therapy, respectively.[1] Until recently, new markers have been tested one, or at most two, at a time, a process that results in inefficient tissue usage, long delays in analysis, and, if positive, implementation.

GENE EXPRESSION PROFILING

Over the last 10 years, technologic advances have resulted in the ability to examine expression of several genes simultaneously. These methods offer the promise of rapid analysis of multiple markers. Perhaps more importantly, they also permit recognition of patterns of gene expression that, in and of themselves, may serve as tumor markers.

Several methods have been developed to look at hundreds and even thousands of gene messages simultaneously. For clinical purposes, the most widely applied of these methods is called gene expression array, which is accomplished by placing a segment of DNA representing a single gene onto a "spot" adhered to a solid medium, such as a glass slide, a "chip," or a fibrous mesh membrane. Using this method, one can place hundreds or even thousands of DNA spots, each representing a single gene, onto a single slide or chip.[2,3] RNA is harvested from the sample of interest, amplified, labeled with a fluorescent dye, and then incubated for hybridization with the complementary DNA sequence contained within the spots on the chip, usually in the presence of RNA from a control sample labeled with a counter-fluorescent dye. The chip/slide is then automatically scanned, and relative levels of fluorescence, representing increased or decreased levels of RNA message for each sample compared to the control, are collected and analyzed using sophisticated statistical techniques. In most cases, the vast majority of genes (so-called "house-keeping" genes) are neither up- nor downregulated compared to control RNA. However, when cells or tissues from one state are compared to another, a small minority (often no more than 100 to 200) of genes are found to be over- or under-expressed by magnitudes of two- to three-fold or more. These expression patterns can be displayed in a format in which the expression of each gene is provided for each tumor, usually based on color differences (Fig 1). For example, genes that are overexpressed for a certain tumor are symbolized by red and those that are underexpressed are symbolized by green. Computer algorithms have been developed that will cluster specimens based simply on similarity of gene expression patterns. Although rarely do two tumors have identical matches of expression of the

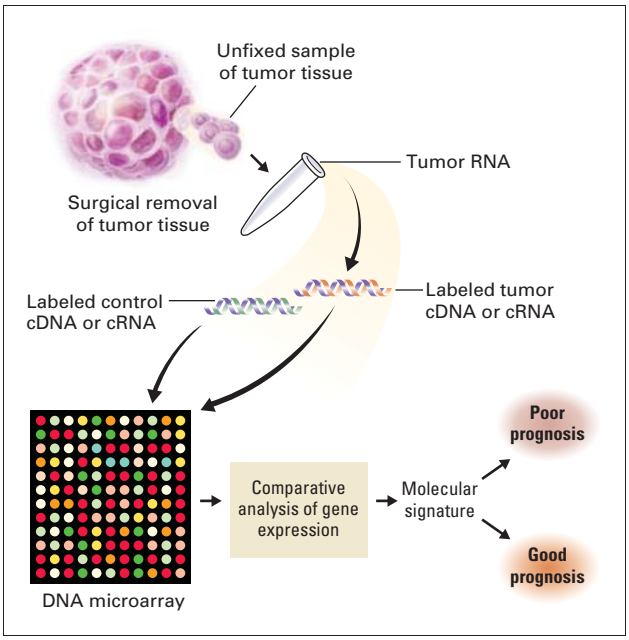

Fig 1. Gene-expression profiling. Unfixed samples of tumor tissue obtained during surgery are the starting material for gene-expression profiling. The expression levels of a set of prognostically relevant genes are determined by DNA-microarray analysis. The resulting molecular signatures allow the patients to be classified into groups with a poor prognosis or a good prognosis, thus facilitating therapeutic decision making. Reprinted with permission from Sauter G, Simon R: Predictive molecular pathology. N Engl J Med 347:1995-1996, 2002. Copyright 2002 Massachusetts Medical Society.

entire gene set; this format permits visual recognition of tumors that share so-called "signatures." Simplistically, this recognition is similar to recognition of the tartan plaids that distinguish one Scottish clan from another. In other words, the details of the individual thread colors are not as important as the overall pattern.

For example, in an early application of this technique, distinct gene expression signatures were identified when normal breast tissue was compared with breast cancer cells.[4] Unstructured cluster analysis has been performed in which the computer places individual tissues into groups (or "clusters") based solely on gene expression patterns, without regard to any preconceived biologic or clinical features. These investigators observed that breast cancer tissues could be grouped into six categories, which they designated basal-like; ErbB-2 positive; normal basal-like; and luminal types A, B, and C.[5,6] In a preliminary study, they observed that luminal A type breast cancers, which are principally positive for hormone receptors and negative for ErbB-2, appeared to have a much more favorable outcome than the other categories.[6]

APPLICATION OF GENE EXPRESSION PROFILING TO DETERMINE CLINICAL OUTCOMES

These early studies have led other investigators to pursue similar approaches to identify signature patterns associated with patient outcomes and prognoses. These preliminary efforts have been remarkably successful, even with very small sample sizes, suggesting the power of signature pattern recognition.[6-9] In a case control study from the Netherlands Cancer Institute (Amsterdam, the Netherlands), one such pattern, consisting of 70 genes, was developed using archived frozen tissue from 117 young, node-negative women with breast cancer for whom 10 or more years of follow-up was available.[9] In this study, tumors from patients who suffered rapid relapses after primary therapy had gene expression profiles that were quite distinct from those who remained disease free. These gene expression profiles were then applied to a second validation set of 295 frozen tissues collected from young women with very similar results (Fig 2).[10] Indeed, it appeared that this 70-gene profiler more accurately predicted outcomes than more classically accepted clinical criteria. In a separate study, investigators from M. D. Anderson Cancer Center (Houston, TX) have suggested that gene expression profiling of cellular material collected by fine needle aspiration before neoadjuvant chemotherapy may identify those women least likely to achieve complete pathologic response.[11] These exciting results suggest that expression profiling can be successfully performed using very small quantities of tissue. Taken together, the data from these preliminary studies suggest that gene expression profiling and development of signature pattern recognition may provide a powerful tool to identify prognosis and likelihood of benefit or resistance to selected therapeutic agents.

DETERMINATION OF CLINICAL UTILITY OF NEW PROGNOSTIC FACTORS

Do these studies provide us with a new tumor marker that should be routinely ordered for newly diagnosed breast cancer patients? Analyses of new factors are best performed in the context of prospective clinical trials in which the prognostic or predictive question can be directly addressed as a function or even an objective of the trial design.[12,13] In this regard, prognostic factors can only be studied in the absence of the therapy one might apply if the patient's prognosis is sufficiently poor to justify it. Therefore, studies of new prognostic factors that do not account for treatment are highly confounded, and conclusions regarding their clinical utility are difficult to determine. Likewise, studies of predictive factors also require rigorous clinical trial design. As in all clinical science, observations are most likely to be valid if they are made in a hypothesis-driven study where the clinical trial design compares patients who received the treatment of interest with a control group who did not, preferably in a prospectively randomized fashion. Such studies are considered to be Class I Levels of Evidence.[14] Unfortunately, most tumor marker studies do not achieve high levels of evidence, and they are more useful to develop new hypotheses rather than confirming preexisting ones.

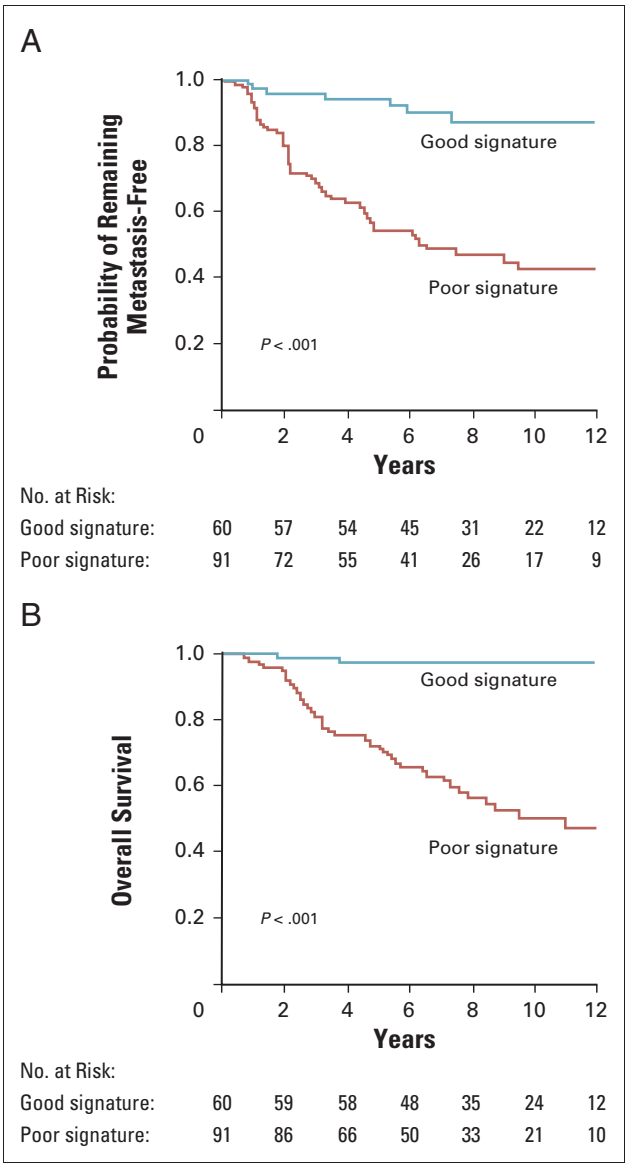

Fig 2. Survival analysis by gene expression profiling for lymph-node–negative breast cancer patients. Kaplan-Meier analysis of the probability that patients would (A) remain free of distant metastases and the probability of (B) overall survival according to whether they had a 70-gene good prognosis or a poor prognosis signature. The *P* values were calculated with use of the log-rank test. Reprinted with permission from van de Vijver MJ, He YD, van 't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999-2009, 2002. Copyright 2002 Massachusetts Medical Society.

APPLICATION OF MULTIPLEX GENE EXPRESSION ARRAY ASSAYS TO FORMALIN-FIXED, PARAFFIN-EMBEDDED TISSUES

The many known and unknown variables that may affect results, such as treatment, are often well controlled in clinical trials conducted by the large cooperative groups in the United States, Europe, and elsewhere. Until recently, gene expression array technology required frozen tissue, and it has been rare that such tissues are available in Cooperative Group archived banks. Most frozen tissue banks have been collected in single institution specimen procurement studies. These banks are often reasonably small and tissues have been collected without regard to patient treatment issues or control of other important variables. Indeed, the European Organisation for the Research and Treatment of Cancer is planning a prospective randomized trial to evaluate the established 70-gene prognostic profile using frozen material in Europe. Nonetheless, for the most part, the large cooperative clinical trials groups have found collection and storage of frozen breast cancer tissue quite difficult, especially with consequent decrease in tumor size associated with widespread screening.[15]

The cooperative groups have, however, already been quite successful in collecting and banking formalin-fixed, paraffin-embedded (FFPE) tissue blocks both retrospectively from completed trials and, more recently, prospectively from ongoing clinical trials.[16] Recently, multiplex, real-time reverse transcriptase polymerase chain reaction (RT-PCR) assays have been developed that permit interrogation of several hundred genes using limited amounts of FFPE sections (Fig 3).[17,18] These assays have been found to accurately quantitate mRNA of known genes, such as ER, even when applied to blocks that have been stored for decades. Therefore, they provide great promise for analysis of the large sample populations that have been studied carefully in prospective clinical trials and for whom long-term outcomes are known.

These technologic advances provide an opportunity to ask hypothesis-driven questions, such as which patients

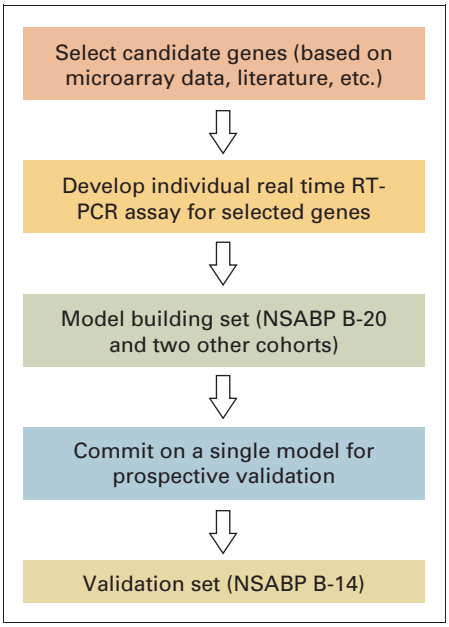

Fig 3. Strategy used for the development of OncoType Dx (Genomics Health Inc, Redwood City, CA) multigene prognostic assay for estrogen receptor–positive, node-negative, tamoxifen-treated breast cancer. RT-PCR, reverse transcriptase polymerase chain reaction; NSABP, National Surgical Adjuvant Breast and Bowel Project.

should receive adjuvant chemotherapy. This question is especially pertinent to those who, by classical prognostic factors, have a favorable prognosis. In general, this question is most germane to node-negative patients with hormone receptor–rich tumors, who will almost all receive adjuvant endocrine treatment. Overall, such patients are generally considered to have a low risk of recurrence: approximately 15% over 5 years.[19] Accumulated data suggest that adjuvant chemotherapy reduces the risk of recurrence and death in this patient group, but obviously the benefit can only be applicable to those will suffer relapse in the absence of therapy.[20,21] Therefore, more accurate prognostic and predictive factors to identify those most likely to relapse and to benefit from this more toxic treatment would be of great value. Because they often have small tumors, identifying banks of frozen breast cancer tissues from this category of patients is even more challenging than for those with larger tumors, further emphasizing the importance of expression profiling using fixed tissues.

The National Surgical Adjuvant Bowel and Breast Project (NSABP) applied multiplex RT-PCR of 185 candidate genes preselected from the previously published gene expression studies to two of their clinical trials in which patients with ER-positive, node-negative breast cancer were treated with tamoxifen.[22] To develop a recurrence score algorithm, these investigators first profiled archived samples that were available from 233 patients (approximately 30%) who participated in clinical trial B-20 and from patients from two other data sets. All patients in NSABP B-20 were treated with tamoxifen and were randomly assigned to chemotherapy or not.[23] In this test set, 21 genes (five of which were control genes) could be used 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 $>$ 30%). Tissue samples from a second NSABP trial, B-14, were used to independently validate these results. In NSABP B-14, which has a median follow-up exceeding 14 years, node-negative, ER-positive patients were assigned to tamoxifen versus placebo.[23] The 21-gene profiler recurrence score algorithm from 668 tissues produced strikingly similar results for prognosis to those they predicted from the B-20 test set.[22] These two studies suggest that this assay may accurately identify up to 50% of patients whose prognosis with hormone therapy alone is so favorable that they would probably forego adjuvant chemotherapy.

VERY RECENT RESULTS AND FUTURE DIRECTIONS

Both the European and the NSABP investigators have recently reported follow-up studies of their respective assays, presented in abstract form only. The TransBIG group has applied the 70-signature profiler to a series of 350 patients from five European and one US center and, again, have observed a dramatic difference in outcomes between those with favorable and unfavorable profiles.[24] The NSABP has reported that the 21-gene profiler assay is also prognostic for the untreated group of women followed in B14. Furthermore, they have observed that it is predictive for benefit from tamoxifen in this study and that it is predictive of benefit from the chemotherapy in B20.[25] The hazard ratio for recurrence for those who received chemotherapy versus those who received tamoxifen alone in NSABP B-20 cohort was 1.3 (95% CI, 0.5 to 3.8) for the low-risk group, 0.6 (95% CI, 0.2 to 1.6) for intermediate-risk group, and 0.3 (95% CI, 0.1 to 0.5) for high-risk group (test for interaction between chemotherapy and the profiler assay, $P = .038$). These data suggest that patients in the low-risk group, which make up about 50% of the node-negative, ER-positive cohort, do not benefit from chemotherapy, whereas those who are at high risk for recurrence appear to derive clinically and statistically significant benefit from chemotherapy.

Should these assays be widely adopted for routine clinical use, especially in node-negative, hormone receptor–positive patients? There are several caveats that make the answer to this question less than a resounding "yes." First, there are some technical concerns. As noted, these assays are based on RNA analysis, and RNA is notoriously unstable. Several factors, including prolonged time from excision to freezing, or fixation and prolonged storage in formalin-fixed paraffin blocks, can produce wide variability in mRNA quality. Indeed, the multiplex RT-PCR assay used by the NSABP is based on identification and quantification of RNA fragments. Thus, careful attention must be given to processing, storage, and preparation techniques. Newer techniques of tissue handling, such as proprietary high-salt fixative solutions (RNAlater; Ambion, Austin, TX), may obviate some of these concerns, but widespread adoption of such buffers in standard pathology suites will require substantial validation that this approach is truly of clinical value.[26] Encouraging results from a pilot Dutch multi-institutional pilot study suggested that good quality RNA can be harvested from material maintained in RNA-later in over 95% of cases. This approach may provide the advantage that full expressed genome evaluation can be performed.

The NSABP was able to harvest and profile mRNA from 99% of the tissue blocks they had available for testing. However, comparison of results between B14 and B20 suggested that some genes, such as *HER-2*, were not as accurately quantitated in B-14 as in B-20; these data teach us the importance of examining multiple markers rather than a handful to maintain the stability of performance in clinical setting. Furthermore, in both the European and NSABP studies, samples were collected for purposes other than the multiple gene expression assays, raising issues of selection bias, even in studies of this magnitude. For example, the NSABP tissues represent less than one half of

those women who participated in the trials. Prospective, multi-institutional, randomized trials of both the frozen tissue and FFPE-based assays are being planned in North America and Europe.

Regardless, it appears that assays that evaluate expression of multiple genes simultaneously, using sophisticated statistical analyses to identify and quantify signature patterns, offer incredibly powerful means to further subgroup patients into those likely or not likely to benefit from adjuvant therapies. These assays might identify patients with such a good prognosis that, regardless of the effectiveness of therapy, they would elect to forego it. On the other hand, as predictive factors, they might identify those patients who are most likely to respond to a specific therapy, much as ER and HER-2 do for endocrine and trastuzumab therapy, respectively. With the ability to study large banks of FFPE tissues, the analysis of large numbers of tissues collected by the cooperative groups now appears to be feasible. Likewise, with new technologies, prospective recognition of the importance of collecting appropriately processed frozen or fixed tissues makes application of gene-expression profiling possible. These advances should permit selective investigations of predictive factors for individual agents that have been tested in such trials, such as the taxanes.[27] In summary, these new assays, with further validation, will almost certainly permit more efficient application of these life-saving, but costly, and often toxic, treatments.

Authors' Disclosures of Potential Conflicts of Interest

The following authors or their immediate family members have indicated a financial interest. No conflict exists for drugs or devices used in a study if they are not being evaluated as part of the investigation. Consultant: Daniel Hayes, Genomics Health Inc. For a detailed description of these categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section of Information for Contributors found in the front of every issue.

REFERENCES

1. Bast RC Jr, Ravdin P, Hayes DF, et al: 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol 19:1865-1878, 2001

2. Khan J, Simon R, Bittner M, et al: Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res 58: 5009-5013, 1998

3. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868, 1998

4. Perou CM, Jeffrey SS, van de Rijn M, et al: Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci U S A 96:9212-9217, 1999

5. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. Nature 406:747-752, 2000

6. Sorlie T, Perou CM, Tibshirani R, et al: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869-10874, 2001

7. Ahr A, Holtrich U, Solbach C, et al: Molecular classification of breast cancer patients by gene expression profiling. J Pathol 195:312-320, 2001

8. West M, Blanchette C, Dressman H, et al: Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci U S A 98:11462-11467, 2001

9. van't Veer LJ, Dai H, van de Vijver MJ, et al: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-536, 2002

10. van de Vijver MJ, He YD, van't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347: 1999-2009, 2002

11. Ayers M, Symmans WF, Stec J, et al: Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol 22: 2284-2293, 2004

12. Hayes DF, Bast R, Desch CE, et al: A tumor marker utility grading system (TMUGS): A framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst 88:1456-1466, 1996

13. Hayes DF, Trock B, Harris A: Assessing the clinical impact of prognostic factors: When is ''statistically significant'' clinically useful? Breast Cancer Res Treat 52:305-319, 1998

14. Hayes DF: Clinical importance of prognostic factors: Moving from scientifically interesting to clinically useful, in Bronchud MH, Foote M, Giaccone G, et al (eds): Principles of Molecular Oncology (2nd Ed). Totowa, NJ, Humana Press Inc, 2004, pp 51-72

15. Cady B: New era in breast cancer. Impact of screening on disease presentation. Surg Oncol Clin N Am 6:195-202, 1997

16. Schilsky RL, Dressler LM, Bucci D, et al: Cooperative group tissue banks as research resources: The Cancer and Leukemia Group B tissue repositories. Clin Cancer Res 8:943-948, 2002

17. Cobleigh M, Bitterman P, Baker J, et al: Tumor gene expressing predicts distant disease-free survival in breast cancer patients with 10 or more positive nodes: High throughput RT-PCR assay of paraffin-embedded tumor tissues. Proc Am Soc Clin Oncol 22:850a, 2003 (abstr 3415)

18. Paik S, Shak S, Tang G, et al: Multi-gene RT-PCR assay for predicting recurrence in node negative breast cancer patients-NSABP B-20 and B-14. Breast Cancer Res Treat 82:S11, 2003 (abstr 16)

19. Ravdin PM, Siminoff LA, Davis GJ, et al: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 19:980-991, 2001

20. Early Breast Cancer Trialist's Collaborative Group: Tamoxifen for early breast cancer: An overview of the randomised trials. Lancet 351: 1451-1467, 1998

21. Fisher B, Dignam J, Tan-Chiu E, et al: Prognosis and treatment of patients with breast tumors of one centimeter or less and negative axillary lymph nodes. J Natl Cancer Inst 93: 112-120, 2001

22. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351: 2817-2826, 2004

23. Fisher B, Jeong JH, Bryant J, et al: Treatment of lymph-node-negative, oestrogen-receptor-positive breast cancer: Long-term findings from National Surgical Adjuvant Breast and Bowel Project randomised clinical trials. Lancet 364:858-868, 2004

24. Piccart MJ, Loi S, van 't Veer LJ, et al: Multi-center external validation study of the Amsterdam 70-gene prognostic signature in node negative untreated breast cancer: Are the results still outperfoming the clinical-pathological criteria? Breast Cancer Res Treat 88:S17, 2004 (abstr 38)

25. Paik S, Shak S, Tang G, et al: Expression of 21 genes in the recurrence score assay and prediction of clinical beefit from tamoxifen in NSABP study B-14 and chemotherapy in NSABP B-20. Breast Cancer Res Treat 88:S15, 2004 (abstr 24)

26. Ellis M, Davis N, Coop A, et al: Development and validation of a method for using breast core needle biopsies for gene expression microarray analyses. Clin Cancer Res 8:1155-1166, 2002

27. Henderson IC, Berry DA, Demetri GD, et al: Improved outcomes from adding sequential paclitaxel but not from escalating doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer. J Clin Oncol 21:976-983, 2003

28. Sauter G, Simon R: Predictive molecular pathology. N Engl J Med 347:1995-1996, 2002