

Prognostic Utility of the 21-Gene Assay in Hormone Receptor–Positive Operable Breast Cancer Compared With Classical Clinicopathologic Features

Lori J. Goldstein, Robert Gray, Sunil Badve, Barrett H. Childs, Carl Yoshizawa, Steve Rowley, Steven Shak, Frederick L. Baehner, Peter M. Ravdin, Nancy E. Davidson, George W. Sledge Jr, Edith A. Perez, Lawrence N. Shulman, Silvana Martino, and Joseph A. Sparano

From the Eastern Cooperative Oncology Group, Boston, MA; Sanofi-aventis, Bridgewater, NJ; Genomic Health, Inc, Redwood City, CA; The University of Texas M.D. Anderson Cancer Center, Houston; Southwest Oncology Group, San Antonio, TX; North Central Cancer Treatment Group, Rochester, MN; Cancer and Leukemia Group B, Chicago, IL.

Submitted September 13, 2007; accepted April 10, 2008; published online ahead of print at www.jco.org on August 4, 2008.

Supported in part by the United States Department of Health and Human Services and the National Institutes of Health (Grants No. CA23318 to the Eastern Cooperative Oncology Group [ECOG] statistical center, CA66636 to the ECOG data management center, CA21115 to the ECOG coordinating center, CA25224 to North Central Cancer Treatment Group, CA32291 to Cancer and Leukemia Group B, and CA32012 to Southwest Oncology Group), and by a grant from Sanofi-aventis.

Presented in part at the 43rd Annual Meeting of the American Society of Clinical Oncology, June 1-5, 2007, Chicago, IL; and at the San Antonio Breast Cancer Symposium, San Antonio, TX, December 13-16, 2007.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

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

Clinical Trials repository link available on JCO.org.

Corresponding author: Joseph A. Sparano, MD, Albert Einstein Cancer Center, Montefiore Medical Center, Weiler Division, 1825 Eastchester Rd, 2 South, Rm 47-48, Bronx, NY 10461; e-mail: jsparano@montefiore.org.

© 2008 by American Society of Clinical Oncology

0732-183X/08/2625-4063/\$20.00

DOI: 10.1200/JCO.2007.14.4501

ABSTRACT

Purpose

Adjuvant! is a standardized validated decision aid that projects outcomes in operable breast cancer based on classical clinicopathologic features and therapy. Genomic classifiers offer the potential to more accurately identify individuals who benefit from chemotherapy than clinicopathologic features.

Patients and Methods

A sample of 465 patients with hormone receptor (HR) –positive breast cancer with zero to three positive axillary nodes who did (n = 99) or did not have recurrence after chemohormonal therapy had tumor tissue evaluated using a 21-gene assay. Histologic grade and HR expression were evaluated locally and in a central laboratory.

Results

Recurrence Score (RS) was a highly significant predictor of recurrence, including node-negative and node-positive disease ($P < .001$ for both) and when adjusted for other clinical variables. RS also predicted recurrence more accurately than clinical variables when integrated by an algorithm modeled after Adjuvant! that was adjusted to 5-year outcomes. The 5-year recurrence rate was only 5% or less for the estimated 46% of patients who have a low RS (< 18).

Conclusion

The 21-gene assay was a more accurate predictor of relapse than standard clinical features for individual patients with HR-positive operable breast cancer treated with chemohormonal therapy and provides information that is complementary to features typically used in anatomic staging, such as tumor size and lymph node involvement. The 21-gene assay may be used to select low-risk patients for abbreviated chemotherapy regimens similar to those used in our study or high-risk patients for more aggressive regimens or clinical trials evaluating novel treatments.

J Clin Oncol 26:4063-4071. © 2008 by American Society of Clinical Oncology

INTRODUCTION

Adjuvant cytotoxic chemotherapy reduces the relative risk of recurrence by approximately 30% in women with operable breast cancer.¹ Other adjuvant systemic therapies that reduce recurrence include endocrine therapy for hormone receptor (HR) –positive disease and trastuzumab for disease that overexpresses the human epidermal growth factor receptor 2 (HER-2) protein. With regard to chemotherapy, anthracyclines are now commonly used because they reduce relative recurrence risk by an additional 10% when combined with alkylating agents, such as cyclophosphamide.¹ More recently, numerous trials have demonstrated a modest advantage when

adding the taxanes paclitaxel or docetaxel to doxorubicin-based chemotherapy (whether used concurrently or sequentially after doxorubicin) and have been referred to as third-generation regimens.²⁻⁶ A recent report indicated that the benefits of taxane-containing adjuvant regimens were driven largely by a significant treatment effect in HR-negative disease, with only a modest treatment effect observed in HR-positive disease.⁵ The lower relapse rates in HR-positive disease observed in contemporary clinical trials have created challenges in evaluating new therapeutic strategies. Extremely large trials that must include thousands of patients are now required to detect improved outcomes for new treatment strategies. Use of standardized validated tools for patient

selection could enhance the likelihood of success of such trials and allow them to be performed more efficiently.

Adjuvant! is a standardized validated instrument that projects patient outcomes based on classical clinicopathologic features and therapy, was designed to predict outcome at 10 years, and was validated for this end point.⁶ Although long-term outcome is generally accepted as fully reflective of the clinical utility of an adjuvant therapy, precise estimation of shorter-term outcomes at 5 years may offer advantages if found to be reflective of longer-term outcomes.

Several studies have demonstrated that multigene molecular markers predict relapse more accurately than clinical criteria in patients with negative or positive axillary nodes and in HR-positive or HR-negative disease.⁷⁻¹¹ Of these signatures, only the 21-gene assay (Oncotype DX Recurrence Score [RS]; Genomic Health, Inc, Redwood City, CA) was developed and validated for patients with HR-positive, node-negative disease treated with endocrine therapy and is performed on routinely processed paraffin-embedded tumor tissue. The purpose of this study was to evaluate the prognostic utility of this 21-gene assay in patients with either node-negative or node-positive HR-positive breast cancer treated with doxorubicin-containing chemotherapy and to determine whether it could more reliably predict outcome at 5 years than standard clinicopathologic features. We also compared the prognostic utility of RS compared with clinicopathologic features and treatment information when integrated by an algorithm similar to but not identical to Adjuvant!, with the algorithm adjusted to 5-year outcomes. The purpose of this comparison was not to determine which method for estimating prognosis was superior, but rather to provide a more stringent test of the prognostic utility of the 21-gene assay. The development of such markers offer the potential to more accurately predict patient benefit and outcomes with specific therapies, to make more informed treatment recommendations, and to enrich for high-risk populations to select specific chemotherapy regimens or clinical trials evaluating more novel treatment strategies.

PATIENTS AND METHODS

Study Population and Treatment

The study used tumor specimens and clinical information from patients enrolled on trial E2197 (ClinicalTrials.gov identifier: NCT00003519), coordinated by the Eastern Cooperative Oncology Group (ECOG). The trial included 2,885 eligible and assessable patients with operable breast carcinoma and one to three positive axillary lymph nodes or negative axillary lymph nodes with the primary tumor measuring at least 1.1 cm in size. Patients were randomly assigned to receive four 3-week cycles of doxorubicin 60 mg/m² plus cyclophosphamide 600 mg/m² or docetaxel 60 mg/m², plus endocrine therapy (if HR-positive) for 5 years or longer. Tamoxifen (20 mg daily for 5 years) was recommended beginning after completion of chemotherapy when the trial was initiated, although approximately 40% of patients eventually took an aromatase inhibitor at some point before or after 5 years when it was shown that these agents were more effective than tamoxifen.¹² The treatment arms were well balanced with regard to median age (51 years), proportion with lymph node-negative disease (65%), and HR-positive disease (64%). After a median follow-up of 76 months, there was no significant difference between arms in disease-free survival (the primary study end point), relapse-free interval, or overall survival in the entire study group and in the population included in this analysis.

Specimen Analysis

Tumor specimens were evaluated for [histologic grade](#) using the modified Bloom-Richardson score by a single pathologist (F.L.B.) using 3- or 4- μ m

tissue sections stained by hematoxylin and eosin¹³ and for [estrogen receptor \(ER\)](#), [progesterone receptor \(PR\)](#), and HER-2 expression by [immunohistochemistry](#) by two pathologists simultaneously (F.L.B., S.B.) using two 1.0-mm [tissue microarray](#) cores using previously described methods summarized in the Appendix (online only).^{14,15} There was good concordance between the local institutional pathology laboratory and central laboratory for HR by immunohistochemistry (90%). All specimens were also analyzed for the Oncotype DX Recurrence Score as previously described.¹⁶

Case and Control Selection, End Points, and Statistical Analyses

The analysis included 465 patients with HR-positive disease, of whom 99 patients had experienced relapse (cases) and 366 patients did not (controls). The methods for case and control selection and statistical analyses are described in the Appendix. This includes a description of how the analysis was weighted to compensate for the unbalanced manner in which the cases were sampled relative to the controls. The primary end point was recurrence-free interval (RFI), defined as the time from trial entry to the first evidence of breast cancer recurrence (which included invasive breast cancer in local, regional, or distant sites, including the ipsilateral breast, but excluded new primary breast cancers in the opposite breast).¹⁷

RESULTS

Characteristics of Study Population and RS Distribution

There were no significant differences in the characteristics of sample included in this analysis compared with the not-in-sample group with HR-positive disease (Table 1), with the exception of estimated menopause distributions ($P = .03$) and the proportion with two to three nodes positive ($P = .05$). The raw numbers of patients in the sampled group and weighted estimates of the percents in the E2197 population with an RS that was low (< 18), intermediate (18 to 30), or high (≥ 31) were 198 patients (46%), 142 patients (30%), and 125 patients (24%), respectively. With regard to HER-2 expression based on central testing, 76 (16% weighted) of 465 patients were positive, including 21 (20% weighted) of the 99 patients who experienced relapse. There was concordance between central and local tumor grade in 249 (57%) of 437 cases in which information regarding local grade was known, a level of concordance that is consistent with that of other reports.¹⁶

Correlation of Recurrence With RS as Categorical or Continuous Variable

Continuous RS, without considering other variables, was a highly significant predictor of recurrence overall and for both the node-negative and node-positive patients ($P < .001$, linear trend test, in each case). When evaluated as a categorical variable, RS was also predictive of an elevated risk of recurrence regardless of nodal status (Fig 1A). The risk of recurrence was elevated in patients with two to three compared with zero or one positive nodes but was not elevated in patients with one positive versus zero positive nodes. For patients with an RS less than 18, approximately 3.3% (95% CI, 2.2% to 5.0%) of patients experienced recurrence within 5 years if there were zero to one positive nodes, and 7.9% (95% CI, 4.3% to 14.1%) experienced recurrence if there were two to three positive nodes. When modeled as a smooth continuous function using splines, there was a direct correlation between RS and recurrence risk up to an RS of approximately 40 ($P < .001$; Fig 1B), even if the analysis was restricted to only patients with HER-2-negative disease ($P < .001$; Fig 1C). Risk of recurrence

Table 1. Distribution of Patient Characteristics in Sample Compared With Not-in-Sample Cohort

Factor	In Sample (n = 465)		Not in Sample (n = 1,503)	
	%	SE	%	SE
Chemotherapy				
AT arm	50.1	1.1	50.2	1.1
AC arm	49.9	1.1	49.8	1.1
Age, years				
≤ 45	23.7	2.0	27.1	1.2
46-65	63.7	2.3	63.0	1.3
> 65	12.7	1.6	9.9	0.8
Menopausal status				
Premenopausal	41.4	2.4	47.6	1.3
Postmenopausal	58.6	2.4	52.4	1.3
Tumor size, cm				
≤ 2.0	53.9	2.4	57.5	1.3
2.1-5.0	42.5	2.4	39.3	1.3
> 5.0	3.6	0.9	3.2	0.5
Axillary nodes				
Negative	56.5	1.1	58.1	1.1
One positive	24.0	1.6	26.1	1.1
Two positive	13.5	1.4	10.5	0.8
Three positive	6.1	1.0	5.3	0.6
Local tumor grade				
Low	12.3	1.6	14.4	0.9
Intermediate	48.6	2.4	48.8	1.3
High	33.2	2.2	29.7	1.2
Unknown	5.9	1.2	7.1	0.7
Central tumor grade				
Low	22.6	2.1	NA	
Intermediate	47.5	2.4	NA	
High	29.9	2.2	NA	
Local HER2/neu*				
Positive	21.9	2.0	23.4	1.1
Negative	44.0	2.4	45.1	1.3
Unknown	34.1	2.3	31.5	1.2
Recurrence and survival				
5-year RFI	90.0	0.7	91.1	0.7
5-year OS	92.3	1.1	94.3	0.6
Median follow-up, years	6.3		6.3	

NOTE. The not-in-sample cohort included all patients with hormone receptor-positive disease (determined locally) not included in this analysis. There were no significant differences in the characteristics of sample patients included in this analysis compared with the not-in-sample group with hormone receptor-positive disease, with the exception of estimated menopause distributions ($P = .03$) and the proportion with two to three nodes positive ($P = .05$).

Abbreviations: AT, doxorubicin plus docetaxel; AC, doxorubicin plus cyclophosphamide; RFI, relapse-free interval; OS, overall survival; NA, not applicable (patients not in sample did not have tumor grade centrally reviewed).

*Local HER2 positive included IHC 2+.

did not increase with an RS greater than 40 in this group of patients treated with chemotherapy, consistent with previous reports suggesting greater benefit from chemotherapy if the RS was high.¹⁸

RS and Recurrence Risk When Adjusted for Clinical Variables

Proportional hazard models for recurrence were performed that included centrally determined tumor grade and HER-2 expression, age, tumor size, and number of positive axillary lymph nodes. RS was evaluated as a continuous linear variable in a manner that was similar

to the B14 validation study; the hazard ratio (HR) for recurrence was calculated relative to an increment of 50 RS units, which is half the range of RS values and thus improves comparability of the HR with the HRs based on the clinical variables. When not including RS, factors associated with a significantly increased recurrence risk included two to three positive axillary nodes (but not one compared with zero positive nodes), young age, and tumor grade. When linear RS was included in the model, only two to three positive nodes and young age were associated with a higher risk, and there was also a strong trend for RS (HR for a 50-point difference in RS = 2.12; 95% CI, 0.97 to 4.65; $P = .06$, linear trend test; Table 2). If the model included locally determined tumor grade, which is more reflective of actual clinical practice, RS was highly associated with recurrence (HR = 3.13; 95% CI, 1.60 to 6.14; $P = .0009$). If the model included only 389 patients with HER-2/*neu*-negative tumors, RS was not predictive whether local or central grading was used ($P =$ not significant, linear trend tests).

Comparison of Outcomes Predicted by the 21-Gene Assay and an Integrator of Clinicopathologic Information

This analysis was performed in four steps. We used an integrator of clinicopathologic information that was modeled after Adjuvant!, but adjusted to 5-year outcomes rather than 10-year outcomes. We will therefore refer to this tool as an integrator rather than Adjuvant!, because use of the latter term implies the use of the Adjuvant! model that is predictive of 10-year outcomes. First, outcomes projected by the integrator were determined for each patient using both local grade and central grade (Fig 2), which demonstrates some discordance in prediction based on the grade used. However, the 5-year recurrence rate of 10% (using central grade) predicted by the integrator was identical to the 10% (SE of mean \pm 0.7%) observed recurrence rate, indicating that the integrator performed well in predicting the average outcome for the entire population.

Second, we evaluated the concordance between prediction made by RS and the integrator using either risk group or risk percentile classification for comparison. For risk group comparisons, patients were classified as being in low, intermediate, or high integrator risk groups, proportionate to the raw distribution of RS risk groups of low (43%), intermediate (31%), and high (27%). For risk percentile classification, both RS and integrator recurrence probabilities were converted to ranks (that is, the values were used to rank the recurrence risk of the patients), and the ranks were then scaled to the interval (0, 100). Put another way, this method is similar to transformation of the recurrence risks to recurrence percentiles. Although use of the ranks puts the integrator and RS on a similar footing, the estimated effects are only relevant to populations with exactly the same distribution of values as in this study. The overall concordance rates between the integrator and RS risk groups were 166 (36%) of 465 when local grade was used and 179 (38%) of 465 when central grade was used, indicating poor concordance between the RS and integrator risk groups. When evaluated by risk percentile, the correlation coefficient was 0.10 when using local grade (Fig 3A) and 0.19 when using central grade (Fig 3B), again demonstrating poor concordance.

Third, we compared the predictive accuracy of RS with the integrator using three different methods, including a [multivariate proportional hazards model](#), [receiver operating characteristic \(ROC\) curves](#),

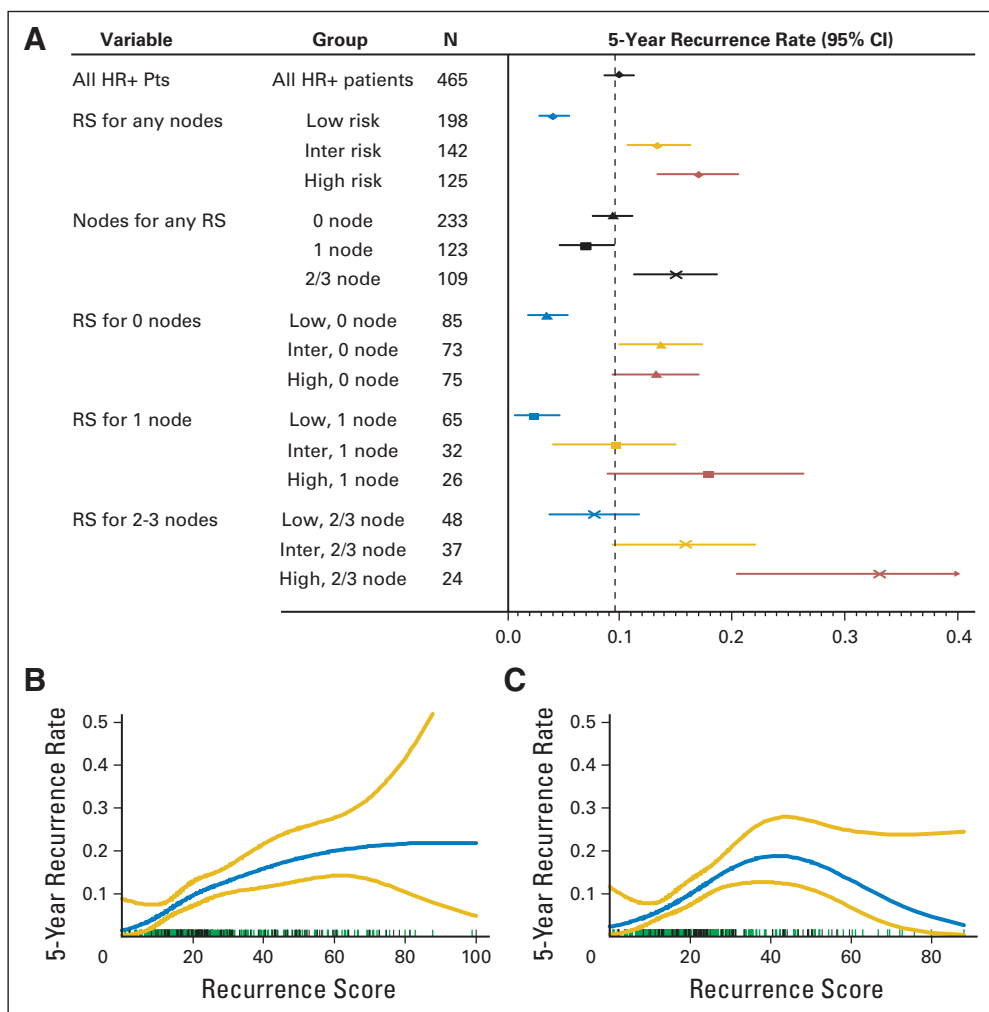


Fig 1. (A) Five-year recurrence rates by Recurrence Score (RS) as a categorical variable by nodal status. RS risk categories: low risk is RS less than 18; intermediate risk is RS 18 to 30; high risk is RS \geq 31. Relationship between RS and recurrence as a continuous variable for (B) all patients with hormone receptor (HR)–positive disease and (C) HR-positive and human epidermal growth factor receptor 2 (HER-2)–negative disease where HER-2–positive disease was excluded. The rug plots above the x-axis show the individual patients with recurrence in black and the individual patients without recurrence in green. The large CIs at higher RS reflect the relatively small number of patients with RS greater than 50. Pts, patients.

and an analysis that evaluated whether RS provided additional information beyond that afforded by the integrator as well as whether the integrator provided additional information that was complementary to RS.

With regard to the proportional hazards model, in a joint linear trend model that included only RS and the integrator risk percentile as variables (and no other factors), the estimated effects for a 50-point difference were HR = 2.64 (95% CI, 1.80 to 3.87; $P < .001$) for RS and

Table 2. Results of Proportional Hazards Models for Recurrence

Ratio	Model Without RS			Model With RS		
	Estimate	95% CI	<i>P</i>	Estimate	95% CI	<i>P</i>
Positive axillary nodes						
1 v 0	0.95	0.57 to 1.57	.002	1.00	0.60 to 1.67	.001
2-3 v 0	2.19	1.42 to 3.36		2.25	1.46 to 3.46	
Tumor size > 2 v \leq 2 cm	1.32	0.84 to 2.09	.23	1.33	0.84 to 2.11	.23
Age, years						
\leq 45 v > 65	2.49	1.09 to 5.71	.008	2.39	1.04 to 5.51	.02
46-65 v > 65	1.23	0.55 to 2.71		1.24	0.56 to 2.75	
HER2, positive v negative	1.17	0.66 to 2.05	.59	0.94	0.53 to 1.68	.84
Grade						
Intermediate v low	1.60	0.81 to 3.15	< .001	1.48	0.75 to 2.94	.20
High v low	3.27	1.64 to 6.50		2.20	0.93 to 5.23	
RS x+50 v x	—	—	—	2.12	0.97 to 4.65	.06

NOTE. Analysis based upon central tumor grade and evaluation of HER2 expression. RS was used as a continuous variable from 0 to 100. Abbreviation: RS, Recurrence Score.

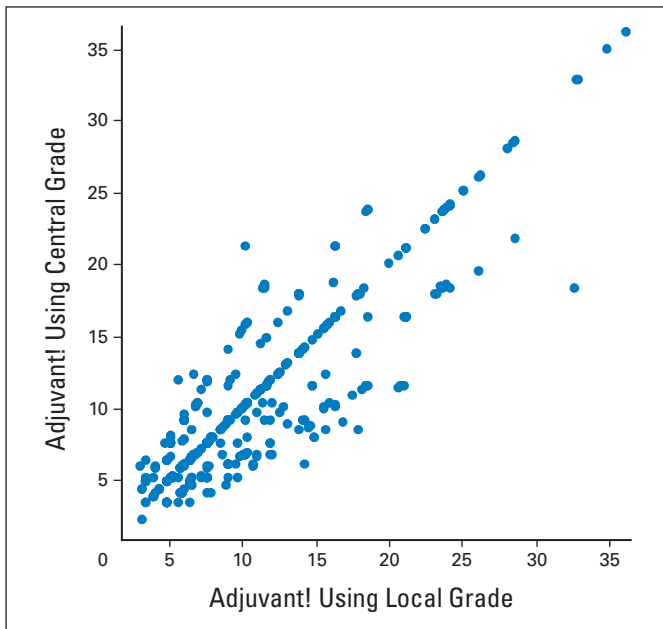


Fig 2. Five-year recurrence probabilities predicted by the integrator using local and central grade. For cases with the same result for local and central grade, the two values are equal (so the points are on the diagonal). Points above the diagonal have a higher risk grade category for central than for local and patients below the diagonal have a higher risk grade category for local than for central.

HR = 1.34 (95% CI, 0.94 to 1.91; $P = .11$) for the integrator using local grade and HR = 2.51 (95% CI, 1.71 to 3.70; $P < .001$) for RS and HR = 1.51 (95% CI, 1.07 to 2.13; $P = .02$) for the integrator using central grade. An RS by integrator interaction term added to either model was not significant ($P = .94$ in the first case, $P = .56$ in the second), showing that the effect of RS is largely independent of the level of the integrator risk.

With regard to ROC curve analysis, ROCs were developed for RS and the integrator, which facilitates evaluation of whether a marker predicts recurrence by a particular time (Fig 4). The area

under the ROC curve (AUC) is a measure of the overall discriminatory power of the marker, with AUC = 0.50 corresponding to no discrimination and AUC = 1.0 corresponding to perfect prediction. The ROC AUC was 0.69 for RS used alone, 0.56 for the integrator using local grade, and 0.61 for the integrator using central grade, indicating that RS performed most accurately in predicting recurrence.

Fourth, we evaluated whether RS provided additional information with regard to the relative risk of recurrence for each integrator risk group, using low RS as the referent and central grade when calculating integrator risk (which optimized the performance of the integrator; Fig 5A). For the 43% of those who were in the low integrator risk group, the risk of recurrence was increased 2.6-fold and 4.0-fold for intermediate and high RS, respectively. For the 30% of patients in the intermediate integrator risk group, the risk of recurrence was increased 9.4-fold and 5.8-fold for intermediate and high RS, respectively. For the 24% of patients in the high integrator risk group, only high RS was associated with a significantly increased risk of relapse (2.6-fold). Therefore, RS provided the greatest discriminatory value in patients estimated to be at a low and intermediate risk by the integrator. We also evaluated whether the integrator provided additional information for each RS risk group (Fig 5B). The relative risk of recurrence was increased 3.15-fold for the high-risk integrator group compared with the low risk group for those who had a low RS and 2.44-fold increased for the intermediate integrator risk group compared with the low risk group for those with an intermediate RS, but not for other groups. This demonstrates that the integrator also provides information that is complementary to RS when the RS is low. Finally, we evaluated the absolute risk of recurrence for those with low, intermediate, and high RS in each integrator risk group (Fig 5C). The risk of recurrence was approximately 3% at 5 years for those with a low RS and were in the low risk integrator groups, identifying a group of patients who might do quite well with a short course of chemotherapy as used in this trial. For each integrator risk group, the risk of recurrence was 10% or higher for those with an intermediate or high RS, identifying patients who may be suitable

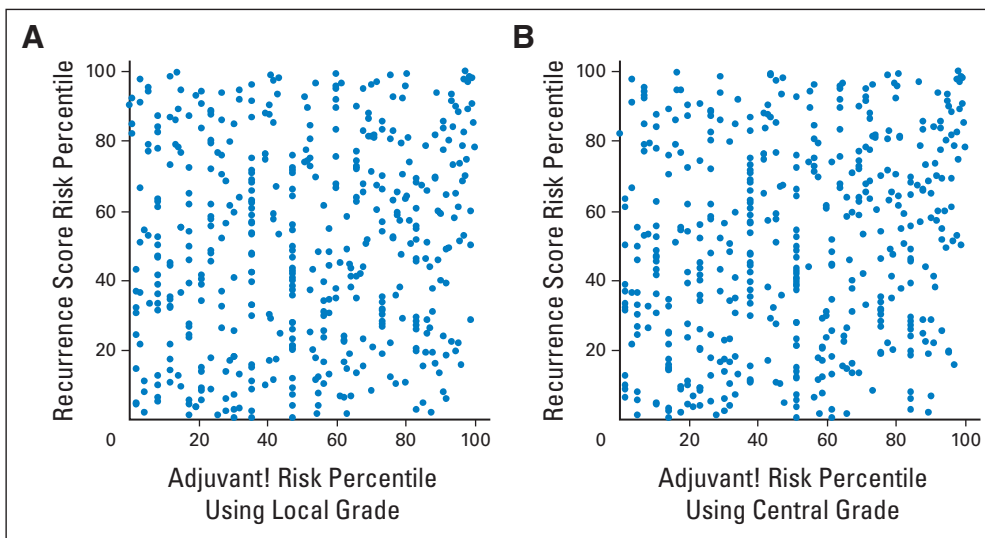


Fig 3. Recurrence score risk percentiles compared with the integrator risk percentiles calculated using (A) local grade (correlation coefficient = 0.10) or (B) central grade (correlation coefficient = 0.19).

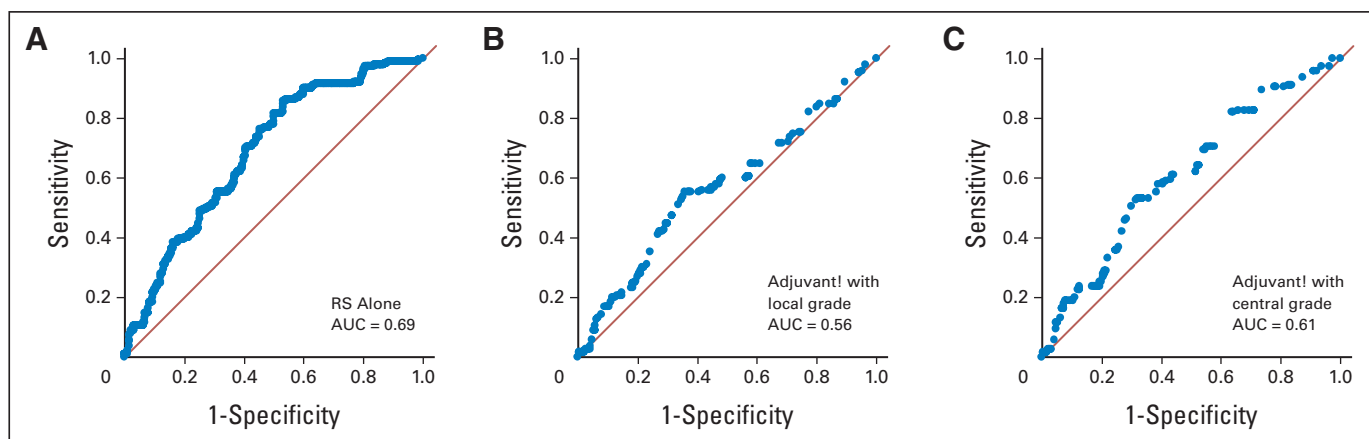


Fig 4. Area under the receiver operating characteristic curves (AUC) for (A) Recurrence Score (RS), (B) the integrator using local grade, and (C) the integrator using central grade.

candidates for more aggressive chemotherapy regimens or clinical trials evaluating novel agents.

DISCUSSION

In the current study of patients with breast cancer with HR-positive disease and zero to three positive nodes who received contemporary chemohormonal therapy, the 21-gene assay was found to predict recurrence more accurately than standard clinicopathologic features individually and when integrated by an algorithm modeled after Adjuvant! that was adjusted to 5-year outcomes. Previous studies have demonstrated that the 21-gene assay predicts distant recurrence more accurately than individual clinical features in patients with HR-positive, axillary lymph node-negative breast cancer treated with tamoxifen.¹⁶ In this study, the assay was shown to provide information on distant and local/regional recurrence risk for individual patients treated with standard chemohormonal therapy independent of and beyond that provided by either nodal status or the integrator, which uses a combination of clinical features. Therefore, the 21-gene assay seems to provide complementary information to classical clinicopathologic features in patients with up to three positive axillary nodes.

For the 46% of patients with a low RS, there was a less than 3% risk of relapse at 5 years if there were zero to one positive nodes and 8% if there were two to three positive nodes. Because there was no arm without chemotherapy treatment, it is not possible here to directly evaluate whether the excellent outcome in those with low RS was related to good prognosis, chemotherapy benefit, or both. The ongoing TAILORx clinical trial (Trial Assigning Individualized Options for treatment) will provide additional information regarding the utility of chemotherapy in patients who have node-negative tumors associated with a midrange RS (11 to 25).¹⁹ Although additional studies will be required to determine whether adjuvant chemotherapy may be withheld from patients with one to three positive axillary lymph nodes with low RS, our findings suggest that this may be possible with additional properly designed studies. A recent report suggests that postmenopausal women with positive axillary lymph nodes and a low RS may not benefit from

adjuvant anthracycline-based chemotherapy when added to tamoxifen,²⁰ providing additional evidence to support this.

The analyses demonstrated a significant correlation between RS and recurrence when analyzed as continuous variable up to an RS of 40. Over its entire range from 0 to 100, RS was of borderline significance when adjusted for centrally determined tumor grade in the proportional hazards model (Table 2), reflecting the greater treatment effect for chemotherapy previously reported for individuals who have a tumor with a high RS.¹⁸ Adjuvant trastuzumab has been shown to reduce the risk of recurrence in HER-2-positive breast cancer,²¹ and HER-2 expression contributes to the RS.¹⁶ When our analysis was restricted to HER-2-negative disease, a significant association between RS and recurrence and pronounced chemotherapy treatment effect for high RS was also observed, indicating that RS provides clinically meaningful information in patients with HER-2-negative disease.

The median follow-up of patients included in this analysis was only 6 years, whereas Adjuvant! was designed for predicting outcomes at 10 years, and both Adjuvant! and the 21-gene profile were validated in populations using data with follow-up exceeding 10 years. The purpose of comparing the 21-gene assay with the integrator modeled after Adjuvant! was to provide a more stringent comparison with classical clinicopathologic factors, as Adjuvant! integrates multiple clinicopathologic and treatment factors. Additional follow-up will be required to determine 10-year outcomes.

Breast cancer mortality rates have declined considerably over the past decade, which has been attributed to improvements in adjuvant therapy and the more widespread use of screening mammography.^{22,23} Moreover, indications for adjuvant chemotherapy have expanded to include even low-risk patients,²⁴ and more effective endocrine therapies have resulted in lower relapse rates for patients with HR-positive disease.²⁵ More recently, adjuvant trastuzumab has been shown to further reduce the risk of recurrence in patients with HER-2-positive breast cancer (irrespective of HR status).²⁶⁻²⁸ Taken together, these advances have led to extremely low rates of relapse for patients diagnosed and treated during the past decade⁵ and have created the need to identify diagnostic tests that may predict outcomes more accurately than clinical criteria. Such diagnostic tests may be used for two purposes: first, to identify

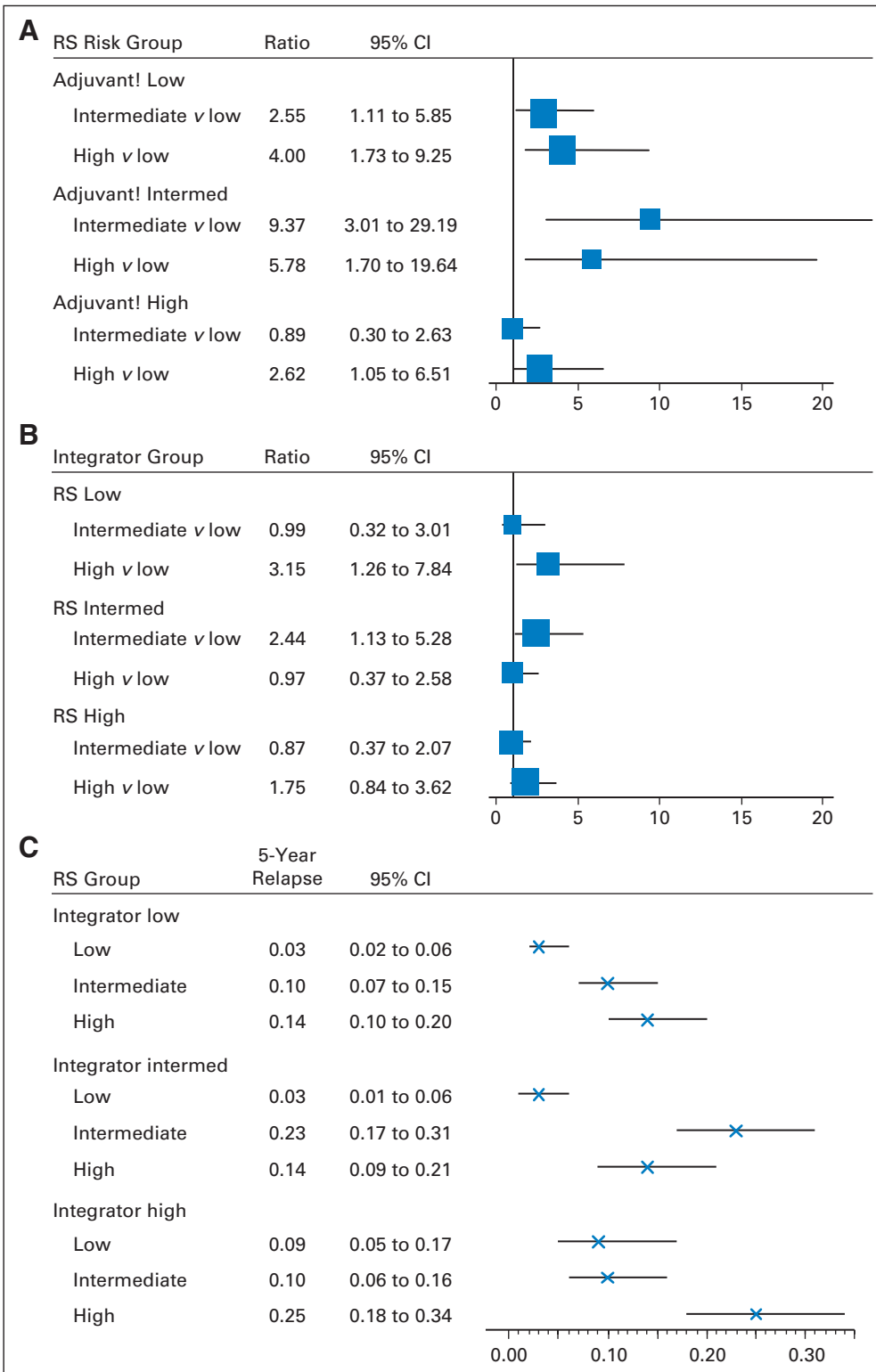


Fig 5. (A) Relative risk of recurrence by Recurrence Score (RS) within subsets defined by the integrator predicted risk level. The area of the box is inversely proportional to the variance of the log hazard ratio, and the line indicates the CI. (B) Relative risk of recurrence by integrator risk group within subsets defined by RS risk groups. The area of the box is inversely proportional to the variance of the log hazard ratio, and the line indicates the CI. (C) Absolute 5-year risk of recurrence (and 95% CIs) by RS risk group within subsets defined by integrator-predicted risk level. Intermed, intermediate.

individuals with HR-positive disease who may be adequately treated with endocrine therapy alone, or endocrine therapy plus a short course of chemotherapy; second, such tests may be used to select high-risk patients for participation in clinical trials, thereby enriching for individuals most likely to benefit from innovative but

potentially more toxic therapies. The TAILORx trial represents the first step toward implementation of this strategy by including patients who meet established clinical criteria for recommending chemotherapy, assigning treatment to hormonal therapy alone if the RS is low (< 11) or chemohormonal therapy if the RS is

elevated (> 25), with random assignment between standard chemohormonal therapy to hormonal therapy alone for those with a midrange RS (11 to 25).¹⁹ Greater use of selection and enrichment strategies that include molecular markers offers potential for providing more informed treatment recommendations, improving the efficiency of clinical trials, and conserving resources.^{29,30}

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure 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 in Information for Contributors.

Employment or Leadership Position: Barrett H. Childs, Sanofi-aventis (C); Carl Yoshizawa, Genomic Health Inc (C); Steve Rowley, Sanofi-aventis (C); Steven Shak, Genomic Health Inc (C); Frederick L. Baehner, Genomic Health Inc (C)

Consultant or Advisory Role: George W. Sledge Jr, Genomic Health Inc (C), Sanofi-aventis (C); Edith A. Perez, Sanofi-aventis (C); Joseph A. Sparano, Sanofi-aventis (C) **Stock Ownership:** Carl Yoshizawa, Genomic Health Inc; Steven Shak, Genomic Health Inc;

Frederick L. Baehner, Genomic Health Inc **Honoraria:** Joseph A. Sparano, Sanofi-aventis **Research Funding:** Joseph A. Sparano, Sanofi-aventis **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Lori J. Goldstein, Robert Gray, Sunil Badve, Steve Shak, Peter M. Ravdin, Nancy E. Davidson, George W. Sledge Jr, Joseph A. Sparano

Financial support: Lori J. Goldstein, Barrett H. Childs, Joseph A. Sparano

Administrative support: Lori J. Goldstein, Joseph A. Sparano

Provision of study materials or patients: Lori J. Goldstein, George W. Sledge Jr, Edith A. Perez, Lawrence N. Shulman, Silvana Martino, Joseph A. Sparano

Collection and assembly of data: Lori J. Goldstein, Robert Gray, Joseph A. Sparano

Data analysis and interpretation: Lori J. Goldstein, Robert Gray, Sunil Badve, Carl Yoshizawa, Steve Rowley, Steven Shak, Frederick L. Baehner, Peter M. Ravdin, Joseph A. Sparano

Manuscript writing: Lori J. Goldstein, Robert Gray, Sunil Badve, Steven Shak, Peter M. Ravdin, Nancy E. Davidson, George W. Sledge Jr, Edith A. Perez, Joseph A. Sparano

Final approval of manuscript: Lori J. Goldstein, Robert Gray, Sunil Badve, Barrett H. Childs, Carl Yoshizawa, Steve Rowley, Steven Shak, Frederick L. Baehner, Peter M. Ravdin, Nancy E. Davidson, George W. Sledge Jr, Edith A. Perez, Lawrence N. Shulman, Silvana Martino, Joseph A. Sparano

REFERENCES

1. Early Breast Cancer Trialists Collaborative Group: Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. *Lancet* 365:1687-1717, 2005
2. 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
3. Mamounas EP, Bryant J, Lembersky B, et al: Paclitaxel after doxorubicin plus cyclophosphamide as adjuvant chemotherapy for node-positive breast cancer: Results from NSABP B-28. *J Clin Oncol* 23:3686-3696, 2005
4. Martin M, Pienkowski T, Mackey J, et al: Adjuvant docetaxel for node-positive breast cancer. *N Engl J Med* 352:2302-2313, 2005
5. Berry DA, Cirincione C, Henderson IC, et al: Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA* 295:1658-1667, 2006
6. Olivetto IA, Bajdik CD, Ravdin PM, et al: Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 23:2716-2725, 2005
7. 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
8. 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
9. Foekens JA, Atkins D, Zhang Y, et al: Multi-center validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *J Clin Oncol* 24:1665-1671, 2006
10. Chang HY, Nuyten DS, Sneddon JB, et al: Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U S A* 102:3738-3743, 2005
11. Ma XJ, Hilsenbeck SG, Wang W, et al: The HOXB13:IL17BR expression index is a prognostic factor in early-stage breast cancer. *J Clin Oncol* 24:4611-4619, 2006
12. Baum M, Buzdar A, Cuzick J, et al: Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early-stage breast cancer: Results of the ATAC (Arimidex, Tamoxifen Alone or in Combination) trial efficacy and safety update analyses. *Cancer* 98:1802-1810, 2003
13. Dalton LW, Page DL, Dupont WD: Histologic grading of breast carcinoma: A reproducibility study. *Cancer* 73:2765-2770, 1994
14. Harvey JM, Clark GM, Osborne CK, et al: Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol* 17:1474-1481, 1999
15. Wolff AC, Hammond ME, Schwartz JN, et al: American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer. *Arch Pathol Lab Med* 131:18, 2007
16. 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
17. Hudis CA, Barlow WE, Costantino JP, et al: Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: The STEEP system. *J Clin Oncol* 25:2127-2132, 2007
18. Paik S, Tang G, Shak S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24:3726-3734, 2006
19. Sparano JA: TAILORx: Trial assigning individualized options for treatment (Rx). *Clin Breast Cancer* 7:347-350, 2006
20. Albain K, Barlow W, Shak S, et al: Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal, node-positive, ER-positive breast cancer (S8814,INT0100). *Breast Cancer Res Treat* 106, 2007 (suppl 1; abstr 10)
21. Viani GA, Afonso SL, Stefano EJ, et al: Adjuvant trastuzumab in the treatment of HER-2-positive early breast cancer: A meta-analysis of published randomized trials. *BMC Cancer* 7:153, 2007
22. Berry DA, Cronin KA, Plevritis SK, et al: Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 353:1784-1792, 2005
23. Shen Y, Yang Y, Inoue LY, et al: Role of detection method in predicting breast cancer survival: Analysis of randomized screening trials. *J Natl Cancer Inst* 97:1195-1203, 2005
24. Carlson RW, Brown E, Burstein HJ, et al: NCCN Task Force Report: Adjuvant Therapy for Breast Cancer. *J Natl Compr Canc Netw* 4:S1-S26, 2006 (suppl)
25. Ryan PD, Goss PE: The evolving role of aromatase inhibitors in the treatment of early-stage breast cancer. *Nat Clin Pract Oncol* 2:596-597, 2005
26. Romond EH, Perez EA, Bryant J, et al: Trastuzumab plus adjuvant chemotherapy for operable

HER2-positive breast cancer. *N Engl J Med* 353:1673-1684, 2005

27. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al: Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med* 354:809-820, 2006

28. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al: Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 353:1659-1672, 2005

29. Sargent DJ, Conley BA, Allegra C, et al: Clinical trial designs for predictive marker validation in

cancer treatment trials. *J Clin Oncol* 23:2020-2027, 2005

30. Jiang W, Freidlin B, Simon R: Biomarker-adaptive threshold design: A procedure for evaluating treatment with possible biomarker-defined subset effect. *J Natl Cancer Inst* 99:1036-1043, 2007

Acknowledgment

We thank Adekunle Raji and other staff members at the Eastern Cooperative Oncology Group Pathology Coordinating Office at the Robert H. Lurie Comprehensive Cancer Center, Chicago, IL.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Glossary Terms

Histologic grade: Histologic grade provides prognostic information in many tumors, including ovarian cancer. It is based on a combination of cellular features (nuclear, cytologic, and architectural). The more nuclear atypia and mitotic figures, the higher the grade.

Immunohistochemistry: The application of antigen-antibody interactions to histochemical techniques. Typically, a tissue section is mounted on a slide and is incubated with antibodies (polyclonal or monoclonal) specific to the antigen (primary reaction). The antigen-antibody signal is then amplified using a second antibody conjugated to a complex of peroxidase-antiperoxidase (PAP), avidin-biotin-peroxidase (ABC) or avidin-biotin alkaline phosphatase. In the presence of substrate and chromogen, the enzyme forms a colored deposit at the sites of antibody-antigen binding. Immunofluorescence is an alternate approach to visualize antigens. In this technique, the primary antigen-antibody signal is amplified using a second antibody conjugated to a fluorochrome. On UV light absorption, the fluorochrome emits its own light at a longer wavelength (fluorescence), thus allowing localization of antibody-antigen complexes.

ROC (receiver operating characteristic) curves: ROC curves plot the true-positive rate (sensitivity) against the false-positive rate (1-specificity) for different cut-off levels of a test. The area under the curve is a measure of the accuracy of the test. An area of 1.0 represents a perfect test (all true positives), whereas an area of 0.5 represents a worthless test.

Tissue microarray: Used to analyze the expression of genes of interest simultaneously in multiple tissue samples, tissue microarrays consist of hundreds of individual tissue samples placed on slides ranging from 2 to 3 mm in diameter. Using conventional histochemical and molecular detection techniques, tissue microarrays are powerful tools to evaluate the expression of genes

of interest in tissue samples. In cancer research, tissue microarrays are used to analyze the frequency of a molecular alteration in different tumor type, to evaluate prognostic markers, and to test potential diagnostic markers.

Multivariate proportional hazards model: Proportional hazards or Cox regression modeling is a general method in medical statistics to analyze the influence of several (patient-specific) covariates on time-to-event end points. No assumption is made concerning the form of the underlying time-to-event curve. The only assumption made is that the effect of the covariates on the hazard rate in the study population is multiplicative and does not change over time.

ER (estrogen receptor): Belonging to the class of nuclear receptors, estrogen receptors are ligand-activated nuclear proteins present in many breast cancer cells that are important in the progression of hormone-dependent cancers. After binding, the receptor-ligand complex activates gene transcription. There are two types of estrogen receptors (α and β). ER α is one of the most important proteins controlling breast cancer function. ER α is present in much lower levels in breast cancer and its function is uncertain. Estrogen-receptor status guides therapeutic decisions in breast cancer.

PgR (progesterone receptor): Like estrogen receptors, progesterone receptors are also nuclear proteins that are activated by the hormone progesterone in breast cancer cells that are hormone-dependent.

HER-2/neu (human epithelial growth factor receptor-2): Also called ErbB2, HER-2/neu belongs to the EGFR family and is overexpressed in several solid tumors. Like EGFR, it is a tyrosine kinase receptor whose activation leads to proliferative signals within the cells. On activation, the HER family of receptors are known to form homodimers and heterodimers, each with a distinct signaling activity. Because HER-2 is the preferred dimerization partner when heterodimers are formed, it is important for signaling through ligands specific for any members of the family. It is typically overexpressed in several epithelial tumors.