

From the Lineberger Comprehensive Cancer Center and Departments of Genetics, Pathology and Laboratory Medicine, and Department of Statistics and Operations Research, Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC; Department of Pathology, University of Utah Health Sciences Center; ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; Genetic Pathology Evaluation Centre, Department of Pathology, Vancouver Coastal Health Research Institute; Departments of Pathology and Radiation Oncology, British Columbia Cancer Agency; Department of Pathology, University of British Columbia, Vancouver, British Columbia, Canada; Genome Sequencing Facility and Division of Oncology, Department of Medicine, Washington University School of Medicine, St Louis, MO; and Department of Pathology, Thomas Jefferson University, Philadelphia, PA.

Submitted May 13, 2008; accepted November 4, 2008; published online ahead of print at www.jco.org on February 9, 2009.

Supported by the Huntsman Cancer Institute/Foundation (P.S.B.), the ARUP Institute for Clinical and Experimental Pathology (P.S.B.), a National Cancer Institute (NCI) Strategic Partnering to Evaluate Cancer Signatures Grant No. U01 CA114722-01 (M.J.E.), an NCI Breast SPORE Grant No. P50-CA58223-09A1 (C.M.P.), a St Louis Affiliate of the Susan G. Komen Foundation CRAFT grant (M.J.E.), and the Breast Cancer Research Foundation (C.M.P. and M.J.E.). Additional support provided by the TRAC facility and Informatics at the Huntsman Cancer Center, supported in part by the NCI Cancer Center Support Grant No. P30 CA42014-19, and the tissue procurement facility at the Alvin J. Siteman Cancer Center at Washington University School of Medicine, which is funded in part by the NCI Cancer Center Support Grant No. P30 CA91842.

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

Corresponding author: Philip S. Bernard, MD, University of Utah, 2000 Circle of Hope, Salt Lake City, UT 84112-5550
e-mail: phil.bernard@hci.utah.edu.

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®).

© 2009 by American Society of Clinical Oncology

0732-183X/09/2708-1160/\$20.00

DOI: 10.1200/JCO.2008.18.1370

Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes

Joel S. Parker, Michael Mullins, Maggie C.U. Cheang, Samuel Leung, David Voduc, Tammi Vickery, Sherri Davies, Christiane Fauron, Xiaping He, Zhiyuan Hu, John F. Quackenbush, Inge J. Stijleman, Juan Palazzo, J.S. Marron, Andrew B. Nobel, Elaine Mardis, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, and Philip S. Bernard

A B S T R A C T

Purpose

To improve on current standards for breast cancer prognosis and prediction of chemotherapy benefit by developing a risk model that incorporates the gene expression-based “intrinsic” subtypes luminal A, luminal B, HER2-enriched, and basal-like.

Methods

A 50-gene subtype predictor was developed using microarray and quantitative reverse transcriptase polymerase chain reaction data from 189 prototype samples. Test sets from 761 patients (no systemic therapy) were evaluated for prognosis, and 133 patients were evaluated for prediction of pathologic complete response (pCR) to a taxane and anthracycline regimen.

Results

The intrinsic subtypes as discrete entities showed prognostic significance ($P = 2.26E-12$) and remained significant in multivariable analyses that incorporated standard parameters (estrogen receptor status, histologic grade, tumor size, and node status). A prognostic model for node-negative breast cancer was built using intrinsic subtype and clinical information. The C-index estimate for the combined model (subtype and tumor size) was a significant improvement on either the clinicopathologic model or subtype model alone. The intrinsic subtype model predicted neoadjuvant chemotherapy efficacy with a negative predictive value for pCR of 97%.

Conclusion

Diagnosis by intrinsic subtype adds significant prognostic and predictive information to standard parameters for patients with breast cancer. The prognostic properties of the continuous risk score will be of value for the management of node-negative breast cancers. The subtypes and risk score can also be used to assess the likelihood of efficacy from neoadjuvant chemotherapy.

J Clin Oncol 27:1160-1167. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Breast cancer is a heterogeneous disease with respect to molecular alterations, cellular composition, and clinical outcome. This diversity creates a challenge in developing tumor classifications that are clinically useful with respect to prognosis or prediction. Gene expression profiling by microarray has given us insight into the complexity of breast tumors and can be used to provide prognostic information beyond standard clinical assessment.¹⁻⁷ For example, the 21-gene OncotypeDx assay (Genome Health Inc, Redwood City, CA) can be used to risk stratify early-stage estrogen receptor (ER) –positive breast cancer.^{4,5} Another strong predictor of outcome in ER-positive disease is proliferation or genomic grade.⁷⁻⁹ In addition, the 70-gene MammaPrint

(Agendia, Huntington Beach, CA) microarray assay has shown prognostic significance in ER-positive and ER-negative early-stage node-negative breast cancer.^{2,3}

The “intrinsic” subtypes luminal A (LumA), luminal B (LumB), HER2-enriched, basal-like, and normal-like have been extensively studied by microarray and hierarchical clustering analysis.^{1,6,10-12} Here, we study the utility of these subtypes alone and as part of a risk of relapse predictor in two cohorts: (1) patients receiving no adjuvant systemic therapy, and (2) patients undergoing paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide (T/FAC) neoadjuvant chemotherapy. The risk of relapse models were compared with standard models using pathologic stage, grade, and routine biomarker status (ER and HER2).

METHODS

Samples and Clinical Data

Patient cohorts for training and test sets consisted of samples with data already in the public domain^{7,13-16} and fresh frozen and formalin-fixed paraffin-embedded (FFPE) tissues collected under institutional review board-approved protocols at the University of British Columbia (Vancouver, British Columbia, Canada), University of North Carolina (Chapel Hill, NC), Thomas Jefferson University (Philadelphia, PA), Washington University (St Louis, MO), and the University of Utah (Salt Lake City, UT). The training set for subtype prediction consisted of 189 breast tumor samples and 29 normal samples from heterogeneously treated patients given the standard of care dictated by their histology, stage, and clinical molecular marker status. The risk of relapse (ROR) models for prognosis in untreated patients were trained using the node-negative, untreated cohort of the Netherlands Cancer Institute (NKI) data set ($n = 141$).¹³ The subtype prediction and ROR models were independently tested for prognosis^{7,14,15} and chemotherapy response.¹⁶ The Hess et al data set¹⁶ used for prediction of chemotherapy sensitivity was not associated with long-term outcome data and was evaluated based on information for pathologic complete response (pCR). Clinical characteristics of the microarray training and test sets are presented in Table 1.¹⁷

Nucleic Acid Extraction

Total RNA was purified from fresh-frozen samples for microarray using the Qiagen RNeasy Midi Kit according to the manufacturer's protocol (Qia-

gen, Valencia, CA). The integrity of the RNA was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The High Pure RNA Paraffin Kit (Roche Applied Science, Indianapolis, IN) was used to extract RNA from FFPE tissues ($2 \times 10\text{-}\mu\text{m}$ sections or 1.5-mm punches) for quantitative reverse transcriptase polymerase chain reaction (qRT-PCR). Contaminating DNA was removed using Turbo DNase (Ambion, Austin, TX). The yield of total RNA was assessed using the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc, Rockland, DE).

Reverse Transcription and qPCR

First-strand cDNA was synthesized using Superscript III reverse transcriptase (first Strand Kit; Invitrogen, Carlsbad, CA) and a mixture of random hexamers and gene-specific primers. PCR amplification and fluorescent melting curve analysis was done on the LightCycler 480 using SYBR Green I Master Mix (Roche Applied Science). A detailed protocol of the PCR conditions can be found in the Appendix (online only).

Microarray

Total RNA isolation, labeling, and hybridizations on Agilent human 1Av2 microarrays or custom-designed Agilent human 22k arrays were performed using the protocol described in Hu et al.⁶ All microarray data have been deposited into the Gene Expression Omnibus¹⁸ under the accession number of GSE10886.

Identification of Prototypic Intrinsic Subtype Samples and Genes

To develop a clinical test that could make an intrinsic subtype diagnosis, we used a method to objectively select prototype samples for training and then predicted subtypes independent of clustering. To identify prototypic tumor samples, we started with an expanded "intrinsic" gene set comprised of genes found in four previous microarray studies.^{1,6,8,11} The normal-like class was represented using true "normals" from reduction mammoplasty or grossly uninvolved tissue, thus we have removed the normal-like class from all outcome analyses and consider this classification as a quality-control measure. A total of 189 breast tumors across 1,906 "intrinsic" genes were analyzed by hierarchical clustering (median centered by feature/gene, Pearson correlation, average linkage),¹⁹ and the sample dendrogram was analyzed using "SigClust".²⁰ A total of 122 breast cancers from 189 individuals profiled by qRT-PCR and microarray had significant clusters representing the "intrinsic" subtypes luminal A (LumA), luminal B (LumB), HER2-enriched, basal-like, and normal-like (Appendix Fig A1, online only). Four additional groups were identified in the training set as significantly reproducible clusters. All four of these groups have similar expression profiles as the luminal tumors and could represent intermediate states or tissue heterogeneity.

Gene Set Reduction Using Prototype Samples and qRT-PCR

A minimized gene set was derived from the prototypic samples using the qRT-PCR data for 161 genes that passed FFPE performance criteria established in Mullins et al.²¹ Several minimization methods were used, including top "N" t test statistics for each group,²² top cluster index scores,²³ and the remaining genes after "shrinkage" of modified t test statistics.²⁴ Cross-validation (random 10% left out in each of 50 cycles) was used to assess the robustness of the minimized gene sets. The "N" t test method was chosen due to having the lowest cross-validation (random 10% left out of each iteration) error. The 50 genes selected and their contribution to distinguishing the different subtypes is provided in Appendix Figure A2 (online only).

Sample Subtype Prediction

The 50 gene set was compared for reproducibility of classification across three centroid-based prediction methods: Prediction Analysis of Microarray (PAM),²⁴ a simple nearest centroid,⁶ and Classification of Nearest Centroid.²⁵ In all cases, the subtype classification is assigned based on the nearest of the five centroids. Because of its reproducibility in subtype classification, the final algorithm consisted of centroids constructed as described for the PAM algorithm²⁴ and distances calculated using Spearman's rank correlation. The centroids of the training set using the 50-gene classifier (henceforth called PAM50) are shown in Appendix Figure A3 (online only).

Table 1. Clinical Characteristics by Cohort

Characteristic	Training Set	No Adjuvant Systemic Therapy*	Neoadjuvant Chemotherapy†
Samples	189	761	133
Median RFS, years	4	9	NA
Age, years			
Mean	58	53	52
Standard deviation	15	13	11
ER			
Positive	114	544	82
Negative	77	195	51
Node			
Positive	96	35	93
Negative	100	710	40
HER2			
Positive	NA	66	33
Negative	NA	192	99
Tumor size, cm			
≤ 2	63	409	13
> 2	136	339	120
Grade			
Low	12	133	2
Medium	56	218	54
High	127	286	75
Subtype			
Luminal A	23	269	37
Luminal B	12	168	27
HER2-enriched	31	120	29
Basal-like	56	128	27
Normal-like	12	76	13

Abbreviations: RFS, relapse-free survival; NA, not applicable; ER, estrogen receptor.

*Compiled from Ivshina et al,¹⁵ Loi et al,⁷ van de Vijver et al,¹³ Wang et al,¹⁴ and University of North Carolina Microarray Database.¹⁷

†Hess et al.¹⁶

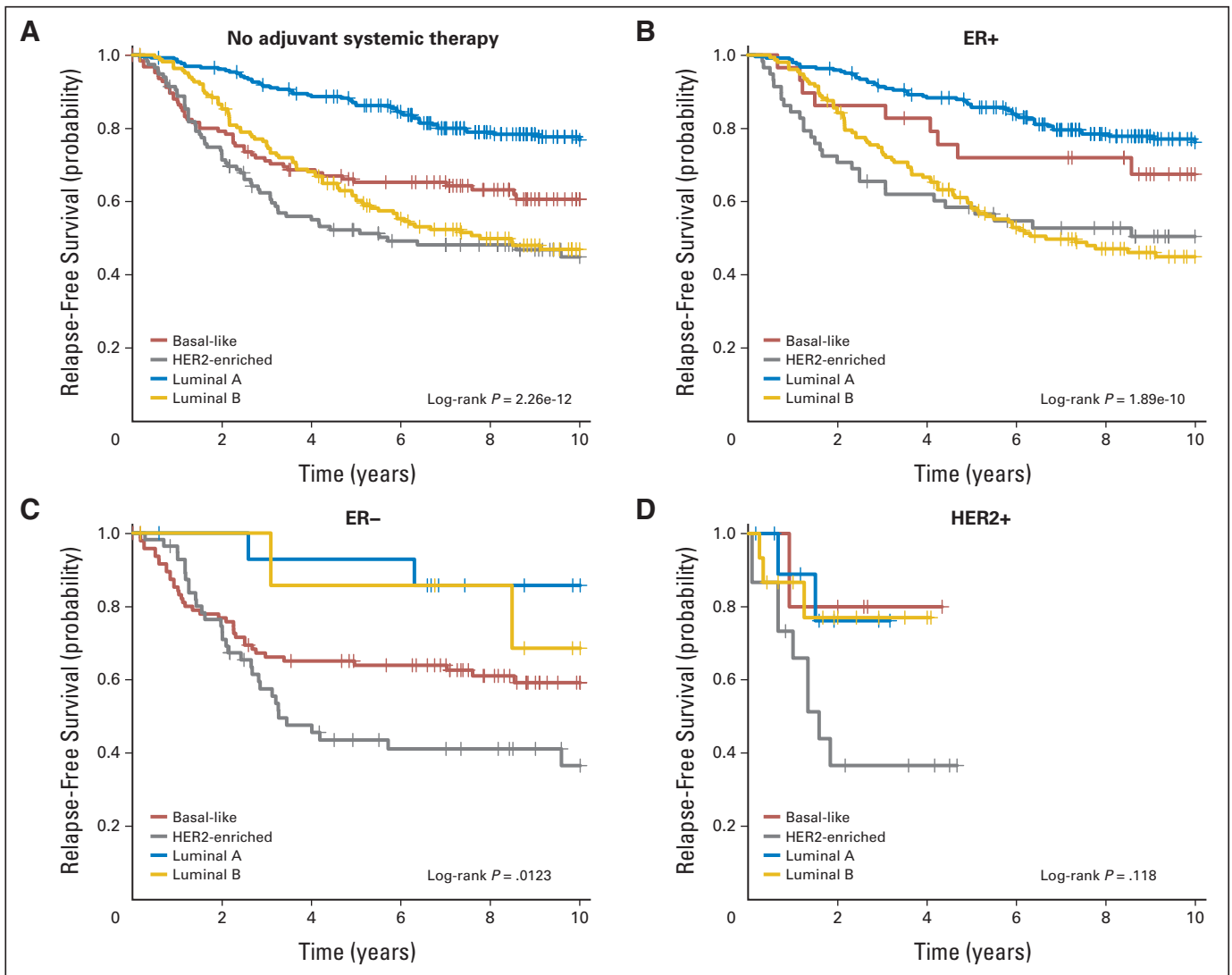


Fig 1. PAM50 intrinsic subtype prognosis for relapse-free survival (RFS). (A) Outcome predictions according to the four tumor subtypes in a test set of 710 node-negative, no systemic adjuvant therapy patients. (B) Outcome by subtype in the subset of patients with estrogen receptor (ER) –positive disease from Figure 1A. (C) Outcome by subtype in patients with ER-negative disease. (D) Outcome by subtype in HER2 $_{clin}$ -positive patients.

Prognostic and Predictive Models Using Clinical and Molecular Subtype Data

Univariate and multivariable analyses were used to determine the significance of the intrinsic subtypes (LumA, LumB, HER2-enriched, and basal-like) in untreated patients and in patients receiving neoadjuvant chemotherapy. For prognosis, subtypes were compared with standard clinical variables (tumor size [T], node status [N], ER status, and histologic grade), with time to relapse (ie, any event) as the end point. Subtypes were compared with grade and molecular markers (ER, progesterone receptor [PR], HER2) for prediction in the neoadjuvant setting because pathologic staging is not applicable. Likelihood ratio tests were done to compare models of available clinical data, subtype data, and combined clinical and molecular variables. Categorical survival analyses were performed using a log-rank test and visualized with Kaplan-Meier plots.

Developing Risk Models With Clinical and Molecular Data

The subtype risk model was trained with a multivariable Cox model using Ridge regression fit to the node-negative, untreated subset of the van de Vijver cohort.¹³ A ROR score was assigned to each test case using correlation to the subtype alone (1) (ROR-S) or using subtype correlation along with tumor size (2) (ROR-C):

$$\text{ROR-S} = 0.05 \cdot \text{basal} + 0.12 \cdot \text{HER2} + \\ -0.34 \cdot \text{LumA} + 0.23 \cdot \text{LumB} \quad (1)$$

$$\text{ROR-C} = 0.05 \cdot \text{basal} + 0.11 \cdot \text{HER2} + \\ -0.23 \cdot \text{LumA} + 0.09 \cdot \text{LumB} + 0.17 \cdot \text{T} \quad (2)$$

The sum of the coefficients from the Cox model is the ROR score for each patient. To classify samples into specific risk groups, we chose thresholds from the training set that required no LumA sample to be in the high-risk group and no basal-like sample to be in the low-risk group. Thresholds were determined from the training set and remained unchanged when evaluating test cases. SiZer analysis was performed to characterize the relationship between the ROR score and relapse-free survival²⁶ (Appendix Fig A4, online only). The 95% CIs for the ROR score are local versions of binomial CIs, with the local sample size computed from a Gaussian kernel density estimator based on the Sheather-Jones choice of window width.²⁷

Comparison of Relapse Prediction Models

Four models were compared for prediction of relapse: (1) a model of clinical variables alone (tumor size, grade, and ER status), (2) ROR-S, (3)

ROR-C, and (4) a model combining subtype, tumor size, and grade. The C-index²⁸ was chosen to compare the strength of the various models. For each model, the C-index was estimated from 100 randomizations of the untreated cohort into two thirds training set and one thirds test set. The C-index was calculated for each test set to form the estimate of each model, and C-index estimates were compared across models using the two sample *t* test.

RESULTS

Distribution of Intrinsic Subtypes in Comparison With Clinical Marker Status

Of the 626 ER-positive tumors analyzed in the microarray test set (Table 1), 73% were luminal (A or B), 11% were HER2-enriched, 5% were basal-like, and 12% were normal-like. Conversely, the ER-negative tumors comprised 11% luminal, 32% HER2-enriched, 50% basal-like, and 7% normal-like. The neoadjuvant study from Hess et al¹⁶ provided an opportunity to analyze the subtype distribution across clinical HER2 (HER2_{clin}) status. Sixty-four percent (21 of 33) of HER2_{clin}-positive were classified as HER2-enriched by gene expression. Only two (6%) of 33 HER2_{clin}-positive tumors were classified as basal-like. Although the majority of the HER2_{clin}-negative tumors were luminal (56%), 9% were classified as HER2-enriched and 24% were basal-like. Thus although the subtype diagnoses have markedly different distributions depending on ER or HER2 status, all subtypes were represented in ER-positive, ER-negative, HER2-positive, and HER2-negative categories. This finding demonstrates that ER and HER2 status alone are not accurate surrogates for true intrinsic subtype status. The intrinsic subtypes showed a significant impact on prognosis for relapse-free survival in untreated (no systemic therapy) patients and when stratified by ER status (Fig 1).

Risk of Relapse Models for Prognosis in Node-Negative Breast Cancer

Cox models were tested using intrinsic subtype alone and together with clinical variables. Table 2 shows the multivariable analyses

of these models in an independent cohort of untreated patients.^{7,13-15} In model A, subtypes, tumor size (T1 v greater), and histologic grade were found to be significant factors for ROR. The great majority of basal-like tumors (95.9%) were found to be medium or high grade, and therefore, in model B, which is an analysis without grade, basal-like becomes significant. Model C shows the significance of the subtypes in the node-negative population. All models that included subtype and clinical variables were significantly better than either clinical alone ($P < .0001$) or subtype alone ($P < .0001$). We trained a relapse classifier to predict outcomes within the context of the intrinsic subtypes and clinical variables. A node-negative, no systemic treatment cohort ($n = 141$) was selected from the van de Vijver microarray data set¹³ to train the ROR model and to select cut-offs (Appendix Fig A5, online only). Figure 2 provides a comparison of the different models using the C-index. There is a clear improvement in prediction with subtype (ROR-S) relative to the model of available clinical variables only (Fig 2A). A combination of clinical variables and subtype (ROR-C) is also a significant improvement over either individual predictor. However, information on grade did not significantly improve the C-index in the combined model, indicating that the prognostic value of grade had been superseded by information provided by the intrinsic subtype model. Figure 2 also presents the use of the ROR-C prognostic model for ROR in a test set of untreated node-negative patients. As was seen on the training data set, only the LumA group contained any low-risk patients (Fig 2B), and the three-class distinction of low, medium, and high risk was prognostic (Fig 2C). Lastly, Figure 2D shows that the ROR-C scores have a linear relationship with probability of relapse at 5 years.

Subtypes and Prediction of Response to Neoadjuvant T/FAC Treatment

The Hess et al¹⁶ study that performed microarray on tumors from patients treated with T/FAC allowed us to investigate the relationship between the subtypes and clinical markers and how each relates to pCR. Table 3 shows the multivariable analyses of the subtypes together with clinical molecular markers (ER, PR, HER2) and

Table 2. Models of Relapse-Free Survival (untreated)

Variable	Model A		Model B		Model C	
	Hazard Ratio	<i>P</i>	Hazard Ratio	<i>P</i>	Hazard Ratio	<i>P</i>
Basal-like*	1.33	.330	1.79	.030	1.58	.066
HER2-enriched*	2.53	.00012	3.25	< .0001	2.90	< .0001
Luminal B*	2.43	< .0001	2.88	< .0001	2.54	< .0001
ER status†	0.83	.38	0.83	.34	0.83	.32
Tumor size‡	1.36	.034	1.43	.012	1.57	.001
Node status§	1.75	.035	1.72	.041	—	—
Histologic grade	1.40	.0042	—	—	—	—
Full v subtype¶		< .0001		< .0001		< .0001
Full v clinical#		< .0001		< .0001		< .0001

Abbreviation: ER, estrogen receptor.

*Luminal A class used as reference state in multivariable analysis.

†Hazard ratios for ER using positive marker in the numerator.

‡Size ≤ 2 cm versus > 2 cm.

§Any positive node.

||Grade encoded as an ordinal variable with three levels.

¶Significant *P* values indicate improved prediction relative to subtype alone.

#Significant *P* values indicate improved prediction relative to clinical data alone.

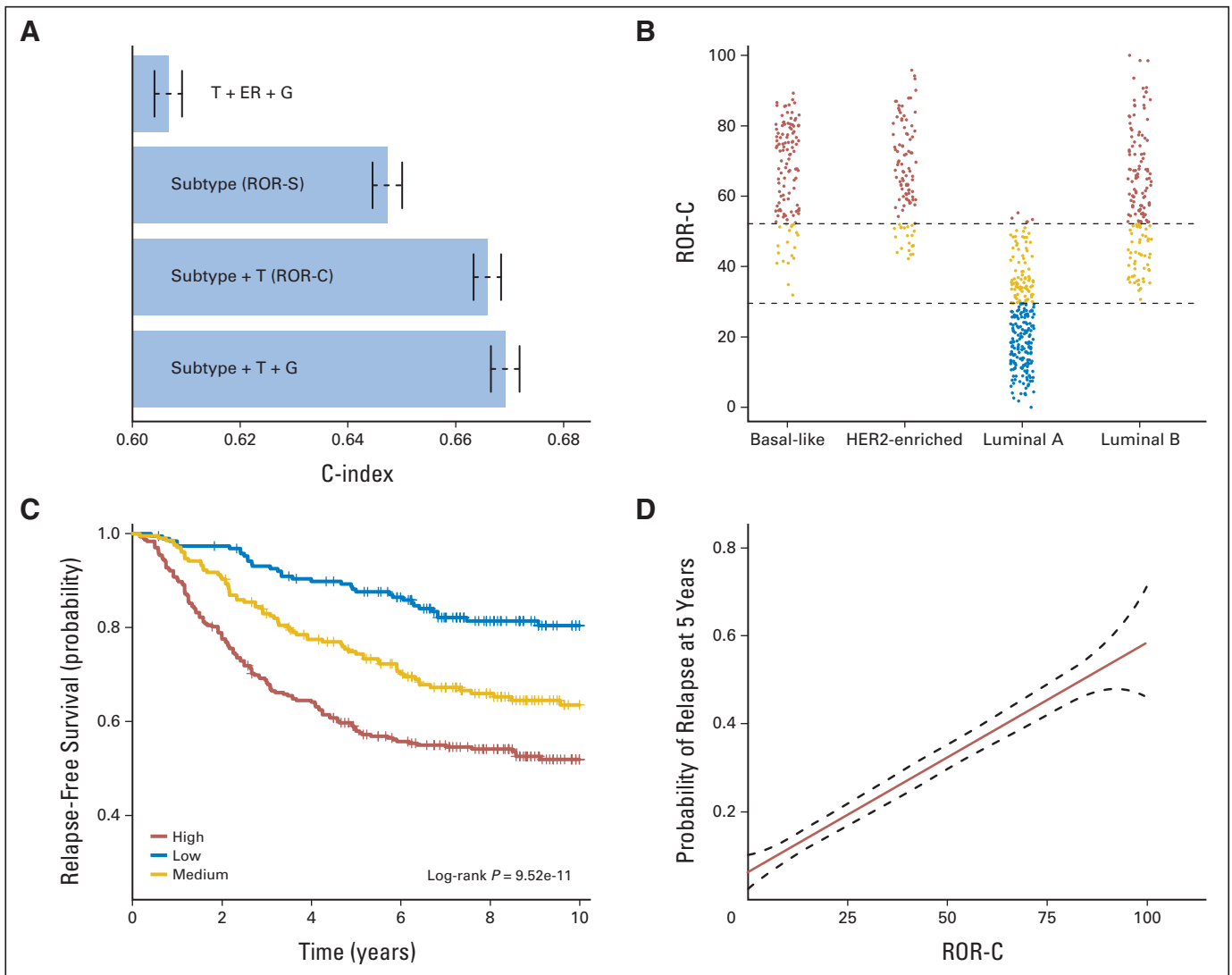


Fig 2. Risk of relapse (ROR) predictions using a test set of node-negative, no systemic therapy patients. (A) C-index analyses of four different Cox models using a test set of node-negative, untreated patients. (B) ROR-C (tumor size and subtype model) scores stratified by subtype. (C) Kaplan-Meier plots of the test set using cut points determined in training. (D) Analysis of the ROR-C model versus probability of survival shows a linear relationship (with the dashed lines showing the 95% CIs). ER, estrogen receptor; RFS, relapse-free survival.

either with (model A) or without (model B) histologic grade. The only significant variables in the context of this study were the intrinsic subtypes. We found 94% sensitivity and 97% negative predictive value for identifying nonresponders to chemotherapy when using the ROR-S model to predict pCR (Fig 3A). The relationship between high-risk scores and a higher probability of pCR (Fig 3B) is consistent with the conclusion that indolent ER-positive tumors (LumA) are less responsive to chemotherapy. However, unlike ROR for prognosis, a plateau seems to be reached for the ROR versus probability of pCR, confirming the presence of significant chemotherapy resistance among the highest risk tumors.

Subtype Prediction and Outcome on FFPE Samples Using qRT-PCR

The subtype classifier and risk predictor were further validated using a heterogeneously treated cohort of 279 patients with FFPE

samples archived between 1976 and 1995. The subtype classifications followed the same survival trends as seen in the microarray data, and the ROR score was significant for long-term relapse predictions (Appendix Fig A6A, online only). This old-age sample set was also scored for standard clinical markers (ER and HER2) by immunohistochemistry (IHC) and compared with the gene expression–based test. Analysis of *ESR1* and *ERBB2* by gene expression showed high sensitivity and specificity as compared with the IHC assay (Appendix Figs A6B and A6C). The advantages of using qRT-PCR versus IHC are that it is less subjective than visual interpretation and it is quantitative.

DISCUSSION

There have been numerous studies that have analyzed interactions between breast cancer intrinsic subtypes and prognosis,^{1,6,11} genetic alterations,²⁹ and drug response.³⁰ Because of the potential clinical

Table 3. Models of Neoadjuvant Response

Variable	Model A		Model B	
	Odds Ratio	<i>P</i>	Odds Ratio	<i>P</i>
Basal-like*	2.96	.021	3.06	.016
HER2-enriched*	1.91	.133	1.99	.117
Luminal B*	2.42	.036	2.43	.035
ER status†	-1.28	.133	-1.31	.124
HER2 status†	0.80	.236	0.85	.208
PR status†	-0.48	.428	-0.46	.446
Histologic grade‡	0.26	.664	—	—
Full v subtype§		.175		.105
Full v clinical		.016773		.0066

Abbreviations: ER, estrogen receptor; PR, progesterone receptor.

*Luminal A class used as reference state in multivariable analysis.

†Hazard ratios for ER, PR, and HER2 are positive marker in the numerator.

‡Grade encoded as an ordinal variable with three levels.

§Significant *P* values indicate improved prediction relative subtype to alone.

||Significant *P* values indicate improved prediction relative to clinical data alone.

value of subtype distinctions, we developed a standardized method of classification using a statistically derived gene and sample set that we have validated across multiple cohorts and platforms. The large and diverse test sets allowed us to evaluate the performance of the PAM50 assay at a population level and in relation to standard molecular markers. An important finding from these analyses is that all of the intrinsic subtypes are present and clinically significant in terms of outcome predictions in cohorts of patients diagnosed with either ER-positive or ER-negative tumors (Fig 1). Thus the molecular subtypes are not simply another method of classification that reflects ER status.

Stratification of the subtypes within *HER2clin*-positive samples did not show significance in outcome predictions; however, there

were fewer numbers and less follow-up in this category. Nevertheless, there was clear separation of the curves for those *HER2clin*-positive patients classified as HER2-enriched (worse prognosis) compared with those with luminal subtypes (better prognosis). We found that 6% of *HER2clin*-positive tumors were classified as basal-like. It has been suggested that *HER2clin*-positive tumors expressing basal markers may have worse outcome when given a chemotherapeutic regimen of trastuzumab and vinorelbine.³¹

Approximately one third of the HER2-enriched expression subtype were not *HER2clin*-positive tumors, suggesting the presence of an ER-negative, nonbasal subtype that is not driven by HER2 gene amplification. The prototype samples selected to represent the HER2-enriched group had high expression of the 17q12-21 amplicon genes (*HER2/ERBB2* and *GRB7*), *FGFR4* (5q35), *TMEM45B* (11q24), and *GPR160* (3q26). In addition, other growth factor receptors such as epidermal growth factor receptor are included within the PAM50 and could potentially also contribute to the HER2-enriched genomic classification.

We found that approximately 10% of breast cancers were classified as normal-like and can be either ER-positive or ER-negative and have an intermediate prognosis. Because the normal-like classification was developed by training on normal breast tissue, we suspect that the normal-like class is mainly an artifact of having a high percentage of normal “contamination” in the tumor specimen. Other explanations include a group of slow-growing basal-like tumors that lack expression of the proliferation genes or a potential new subtype that has been referred to as claudin-low tumors.³² Detailed histologic, immunohistochemical, and additional gene expression analyses of these cases are needed to resolve these issues. Because of the uncertainties, however, the normal-like samples were removed when modeling ROR.

The multivariable analysis for prognosis (ie, no systemic treatment) suggested that the best model was to use subtype with pathologic staging. Because pathologic staging is not available at diagnosis in

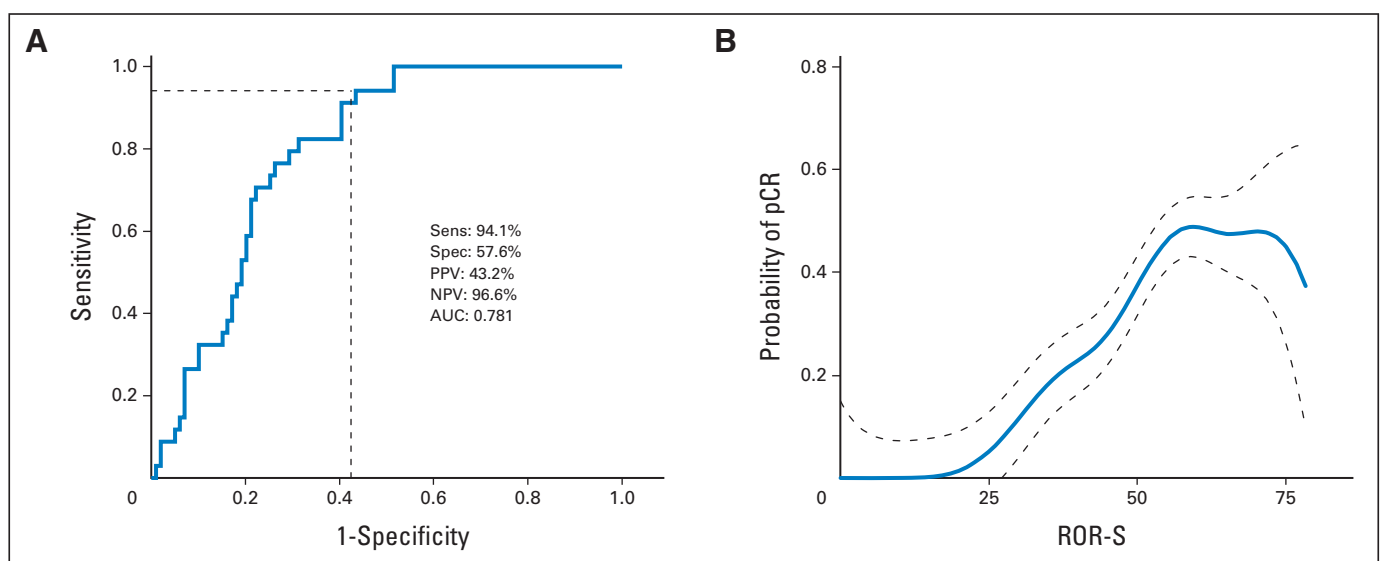


Fig 3. Relationship between risk of relapse (ROR) score and paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide neoadjuvant response. (A) Receiver operating characteristic curve analysis of ROR-S (subtype only model) versus pathologic complete response (pCR) in the Hess et al¹⁶ test set. (B) ROR-S score versus probability of pCR. PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

the neoadjuvant setting, we used histologic grade and clinical biomarkers as the standard for prediction of chemotherapy response before resection. In this context, only the subtypes LumB and basal-like were predictive in the multivariable analysis that included histologic grade, ER, PR, and HER2 status (note that the Hess et al¹⁶ study did not incorporate trastuzumab into the regimen). The ROR score from the subtype-alone model was also the most predictive of neoadjuvant response. One of the major benefits of the ROR predictor is the identification of patients in the LumA group who are at a low ROR on the basis of pure prognosis and for whom the benefit from neoadjuvant therapy is unlikely. Thus the ROR predictor based on subtypes provides similar information as the OncotypeDx Recurrence Score for patients with ER-positive, node-negative disease.^{4,5} Furthermore, the PAM50 assay provides a ROR score for all patients, including those with ER-negative disease, and is highly predictive of neoadjuvant response when considering all patients.

In summary, the intrinsic subtype and risk predictors based on the PAM50 gene set added significant prognostic and predictive value to pathologic staging, histologic grade, and standard clinical molecular markers. The qRT-PCR assay can be performed using archived breast tissues, which will be useful for retrospective studies and prospective clinical trials.

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: Philip S. Bernard, University Genomics Inc (U) **Consultant or Advisory Role:** Philip S. Bernard, University Genomics Inc (U) **Stock Ownership:** Matthew J. Ellis, University Genomics Inc; Charles M. Perou, University Genomics Inc; Philip S. Bernard, University Genomics Inc **Honoraria:** None **Research Funding:** None **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

Provision of study materials or patients: Juan Palazzo, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

Collection and assembly of data: Joel S. Parker, Michael Mullins, Maggie C.U. Cheang, Samuel Leung, David Voduc, Tammi Vickery, Sherri Davies, Christiane Fauron, Xiaping He, Zhiyuan Hu, John F. Quackenbush, Inge J. Stijleman, Juan Palazzo, J.S. Marron, Andrew B. Nobel, Elaine Mardis, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

Data analysis and interpretation: Joel S. Parker, Michael Mullins, Maggie C.U. Cheang, Samuel Leung, David Voduc, Sherri Davies, Christiane Fauron, Xiaping He, Zhiyuan Hu, J.S. Marron, Andrew B. Nobel, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

Manuscript writing: Joel S. Parker, David Voduc, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

Final approval of manuscript: Joel S. Parker, Michael Mullins, Maggie C.U. Cheang, Samuel Leung, David Voduc, Tammi Vickery, Sherri Davies, Christiane Fauron, Xiaping He, Zhiyuan Hu, John F. Quackenbush, Inge J. Stijleman, Juan Palazzo, J.S. Marron, Andrew B. Nobel, Elaine Mardis, Torsten O. Nielsen, Matthew J. Ellis, Charles M. Perou, Philip S. Bernard

REFERENCES

- Sørlie 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
- 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
- van't Veer LJ, Paik S, Hayes DF: Gene expression profiling of breast cancer: A new tumor marker. *J Clin Oncol* 23:1631-1635, 2005
- 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
- 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
- Hu Z, Fan C, Oh DS, et al: The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 7:96, 2006
- Loi S, Haibe-Kains B, Desmedt C, et al: Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 25:1239-1246, 2007
- Perreard L, Fan C, Quackenbush JF, et al: Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res* 8:R23, 2006
- Sotiriou C, Wirapati P, Loi S, et al: Gene expression profiling in breast cancer: Understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 98:262-272, 2006
- Perou CM, Sørlie T, Eisen MB, et al: Molecular portraits of human breast tumours. *Nature* 406:747-752, 2000
- Sørlie T, Tibshirani R, Parker J, et al: Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 100:8418-8423, 2003
- Fan C, Oh DS, Wessels L, et al: Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 355:560-569, 2006
- 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
- Wang Y, Klijn JG, Zhang Y, et al: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365:671-679, 2005
- Ivshina AV, George J, Senko O, et al: Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* 66:10292-10301, 2006
- Hess KR, Anderson K, Symmans WF, et al: Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 24:4236-4244, 2006
- University of North Carolina Microarray Database: GEO Data Sets for Breast Cancer Research Published Papers (Clinical Data updated on 11-06-2007 for Data I, 4-7-2008 for Data II). <https://genome.unc.edu/pubsup/breastGEO/>
- National Center for Biotechnology Information: Gene expression omnibus. <http://www.ncbi.nlm.nih.gov/geo/>
- 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
- Yufeng L, Hayes DL, Nobel A, et al: Statistical significance of clustering for high dimension low sample size data. *J Am Stat Assoc* 103:1281-1293, 2008
- Mullins M, Perreard L, Quackenbush JF, et al: Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues. *Clin Chem* 53:1273-1279, 2007
- Storey JD, Tibshirani R: Statistical methods for identifying differentially expressed genes in DNA microarrays. *Methods Mol Biol* 224:149-157, 2003
- Dudoit S, Fridlyand J: A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biol* 3:RESEARCH0036, 2002

24. Tibshirani R, Hastie T, Narasimhan B, et al: Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 99:6567-6572, 2002
25. Dabney AR: Classification of microarrays to nearest centroids. *Bioinformatics* 21:4148-4154, 2005
26. Chaudhuri P, Marron JS: SiZer for exploration of structures in curves. *J Am Stat Assoc* 94:807-823, 1999
27. Sheather SJ, Jones MC: A reliable data-based bandwidth selection method for kernel density estimation. *J R Stat Soc* 53:683-690, 1991
28. The Design Library: C-index. <http://lib.stat.cmu.edu/S/Harrell/Design.html>
29. Neve RM, Chin K, Fridlyand J, et al: A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10:515-527, 2006
30. Rouzier R, Pusztai L, Delaloge S, et al: Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. *J Clin Oncol* 23:8331-8339, 2005
31. Harris LN, You F, Schnitt SJ, et al: Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. *Clin Cancer Res* 13:1198-1207, 2007
32. Herschkowitz JI, Simin K, Weigman VJ, et al: Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* 8:R76, 2007

