Individualized Survival and Treatment Response Predictions in Breast Cancer Patients: Involvements of Phospho-EGFR and Phospho-Her2/Neu Proteins

Lan Guo*,[1], Jame Abraham[2], Daniel C. Flynn[3], Vincent Castranova[4], Xianglin Shi[4] and Yong Qian*,[4]

[1]MBR Cancer Center/Department of Community Medicine, School of Medicine, West Virginia University, Morgantown, WV 26506-9300, USA

[2]Department of Medicine and Division of Hematology/Oncology, West Virginia University, WV 26506-9300, USA

[3]MBR Cancer Center/Department of Microbiology, Immunology, and Cell Biology, School of Medicine, West Virginia University, Morgantown, WV 26506-9300, USA

[4]The Pathology and Physiology Research Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA

Abstract: Our robust prediction system for individual breast cancer patients combines three well-known machine-learning classifiers to provide stable and accurate clinical outcome prediction (N=269). The average performance of the selected classifiers is used as the evaluation criterion in breast cancer outcome predictions. A profile (incorporating histology, lymph node status, tumor grade, tumor stage, ER, PR, Her2/neu, patient's age and smoking status) generated over 95% accuracy in individualized disease-free survival and treatment response predictions. Furthermore, our analysis demonstrated that the measurement of phospho-EGFR and phospho-Her2/neu is more powerful in breast cancer survival prediction than that of total EGFR and total Her2/neu ($p < 0.05$). The incorporation of hormone receptor status, Her2/neu, patient's age and smoking status into the traditional pathologic markers creates a powerful standard to perform individualized survival and treatment outcome predictions for breast cancer patients.

Keywords: Prognosis, antibody, breast cancer, EGFR, Her2/neu, molecular pathogenesis.

INTRODUCTION

Breast cancer is the most common cancer among women in the United States. There was an estimated 212,600 new cases diagnosed as breast cancer in the United States in 2003. Among them, 39,800 women were expected to die of metastatic diseases [1]. For the past decade, the overall risk of mortality due to breast cancer has been declining with the development of advanced therapies as well as the early detection [2]. However, the survival rate has not been substantially improved for patients with recurrent or metastatic breast cancer [1]. One of the main obstacles to improve the survival rate is to accurately classify breast cancer patients into subgroups of good prognosis and poor prognosis, reflecting a different probability of disease recurrence and survival after therapy. Following this, another unsolved problem is to decide which combination of treatment is most suitable for each individual patient. Accurate survival and treatment response prediction for a given breast cancer patient will guide the selection of the optimal therapy.

Current research in breast cancer outcome prediction based on the pathologic information is not optimized. In addition, efforts have been focused on predictions for populations, rather than for individuals. Most predictions produce a relative chance of survival for an individual based on the experience with a historical population of patients, as approximated by a single Kaplan-Meier survival curve. These methods generally lead to weak predictive power for individual patients. The NIH has defined the statistical guidelines for assessing classification results [3]. Specifically, statistics should be based on false (1 - specificity) and true positive (sensitivity) rates, instead of odds ratios or relative risks. In addition, to avoid over fitting, the developed marker combination in a training sample should be evaluated in a random test sample from the same study [3]. In this study, we constructed a classification system combining several well-known machine-learning algorithms. We evaluated the prediction results according to the new NIH statistics guidelines [3]. This system provides stable and accurate prediction results for individual patients. Using the proposed classification system, the overall accuracy of individualized survival and treatment response predictions is over 95% in a cohort of 269 breast cancer patients.

Substantial efforts have been made to establish the predictive factors for patients with breast cancer during last two decades. The traditional predictive factors are lymph node status, tumor size, histologic type, histologic grade, lym-

*Address correspondence to these authors at the MBR Cancer Center/Department of Community Medicine, School of Medicine, West Virginia University, Morgantown, WV 26506-9300, USA; Fax: (304) 293-4667; Tel: (304) 293-6455; E-mail: lguo@hsc.wvu.edu

Pathology and Physiology Research Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA; Fax: (304) 285-5938; Tel: (304) 285-6286; E-mail: yaq2@cdc.gov

phatic vessel invasion, and hormone receptor status [4]. With the development of molecular biology and cell biology, many new predictive factors have been created, including markers of regulating cell cycle and cell death, Her2/neu, markers of metastasis or metastatic process, lymph node micrometastases, bone marrow micrometastases, and markers of angiogenesis [5]. Although the traditional predictive factors lack the information about the biological diversity of breast cancer and have not reflected the complexity of molecular mechanisms of these diseases, they are still the most valuable criteria for clinicians to decide the relevant therapies [6]. For instance, Adjuvant! (www.adjuvantonline.com) is a prognostic system based on traditional pathological features, including age, ER expression, and grade. It has been independently validated as a reliable aid to clinical decision-making on average breast cancer patients [7]. The newer factors have not yet gained full recognition in the clinic due to their contradictory results and clinical confusion [4,5]. Thus, there is a pressing clinical need to develop a comprehensive profile which includes both traditional and molecular predictive factors to assist clinicians to assess individual breast cancer patient outcomes and to select relevant therapies.

Although lymph node status, tumor size, histologic type, histologic grade, and tumor stage are the most powerful traditional pathologic prognostic and predictive factors for breast cancer patients, the methods for measurements of the traditional pathologic factors have been difficult to standardize [4]. The difference in the measurements could lead to the misclassification of some breast cancer patients. Even under the same standards, the inter-observer variability still occurs frequently among different examiners. Therefore, the traditional factors are inclined to be subjective. To optimize the measurement parameter profile for the individualized survival and treatment response predictions, we explored the efficiency of estrogen receptor (ER), progesterone receptor (PR), and Her2/neu and their association with other objective parameters. The measurement of ER, PR, and Her2/neu is mainly through machine-based analysis, which is objective and amenable to standardization. Our results showed that the sole utilization of these objective factors was unable to achieve highly accurate predictions. In order to provide a comprehensive profile for individualized survival and treatment response predictions in patients with breast cancer, the objective factors were incorporated into the traditional pathologic factors. Our results showed that the combined parameters increased the accuracy of the disease-free survival prediction ($p < 0.000001$) and treatment response prediction ($p < 0.14$) for individual breast cancer patients.

Recently, it has been reported that predictions based on total protein expression levels of both EGFR and Her2/neu may not be sufficient to reflect the functionality of these proteins *in vivo* [8,9]. Both EGFR and Her2/neu are members of receptor tyrosine kinases. They affect cell and tissue functions through the induction of signaling transduction pathways from the outside of a cell to its nucleus through protein phosphorylation and deposphorylation. The critical step of these processes is the phosphorylation of tyrosine residues, which reflects the kinase activities for EGFR and Her2/neu. The phosphorylated EGFR and Her2/neu create the binding sites for the downstream signaling proteins. Some of these downstream signaling proteins are enzymes, which are acti-

vated upon binding to tyrosine-phosporylated EGFR and Her2/neu, and lead to tumorgenesis. Therefore, overexpression of EGFR and Her2/neu does not necessarily indicate that they are functionally activated. This may explain conflicting results concerning the correlation between overexpression of EGFR and breast cancer prognosis and prediction [8]. The difference between the overexpression and the activation of Her2/neu *in vivo* may also reveal the mechanisms which explain why some Her2/neu overexpressed breast cancer patients have no response to Her2/neu-targeted trastuzumab therapy. One way to measure the kinase activities of EGFR and Her2/neu is to use anti-phospho-antibodies to detect tyrosine-phosphorylated EGFR and Her2/neu. In this study, we sought to explore the predictive value of activated EGFR and Her2/neu in the survival prediction for individual breast cancer patients.

MATERIAL AND METHODS

Data Sources

Data for this analysis were obtained from the Biorepository of Clinomics Biosciences Inc, which is a large collection of highly characterized human tissue samples. These samples span a wide range of common diseases, including many forms and stages of cancer, neurological disorders and heart disease. Clinomics has pioneered the development of an emerging new technology known as *Tissue Microarray* to enable researchers to simultaneously study hundreds of individual tissue samples in parallel, establishing the relative levels of protein expression in those samples and allowing conclusions as to the relevance of these proteins to disease to be made [10,11].

Immunohistochemistry

Immunohistochemistry methods were described previously [12,13]. Briefly, tissue specimens were treated with 3% H_2O_2 to quench endogenous peroxidase activity, followed by washing with PBS. After washing, the tissue specimens were first incubated with a specific primary antibody and then with a biotinylated secondary antibody. Substrate-Chromogen was applied to the specimens according to manufacturer's instructions, followed by stained with hematoxylin. The stain was semiquantitatively examined by pathologists (Clinomics BioSciences, Inc.) using the Allred 8-unit system [14]. Staining was scored on a 0 - 5 scale, with 0=no staining. Grades of 1 to 5 represent increased intensity of staining with 5 being strong, dark brown staining. For each tumor, represented by one slide, the tumor epithelial cells proportion score and intensity score were determined. Peritumoral inflammatory and stromal cells were not included in the evaluation. The proportion score included the fraction of positively stained tumor cells and was as follows: 0 = none, 1 = <1/100th; 2 = 1/100th to 1/10th; 3 = 1/10th to 1/3; 4 = 1/3 to 2/3; 5 = >2/3. The estimated average staining intensity of the positive tumor cells was expressed as follows: 0 = none; 1 = weak; 2 = intermediate; 3 = strong [14]. Each protein was measured with six parameters: Cytoplasmic % Intensity defines percent intensity of stain within cytoplasm; Cytoplasmic % Positive defines percent of all cells positive within cytoplasm; Cytoplasmic Total Score defines the product of Cytoplasmic % Intensity and Cytoplasmic % Positive; Nuclear % Intensity defines percent intensity of stain within nucleus; Nuclear % positive defines percent of

Table 1.   Description of the clinical samples (*N* = 269)

Clinical Information	Value	Occurrences
Histology	Infiltrating Ductal Carcinoma Lobular Carcinoma Medullary Carcinoma Papillary Carcinoma Scirrhous Invasive Ductal Carcinoma Tubular Carcinoma	214 17 11 8 13 6
Surgery Procedure	Lumpectomy + node dissection Mastectomy + node dissection	153 116
Age	21 – 50 51-95	96 184
Stage (AJCC) [32]	1 2 3 4	67 121 48 33
Chemotherapy	CMF None	202 67
Radiation	Breast None	153 116
ER	Negative Positive	91 178
PR	Negative Positive	80 189
Her2/neu	Negative Positive	226 43
Metastasis Site	Non-axillary Lymph node Liver Lung Bone Brain None	18 7 4 3 1 236
Smoking	No Yes (> 20 packs/year) Yes (2 packs/day) Yes (3 packs/day)	180 45 7 37
pT (AJCC) [32]	1 2 3 4	80 123 44 22
pN (AJCC)(32)	0 1 2	88 170 11
Nodes Positive (Pathological)	0 1 – 3 4 – 9 $\geq$ 10	72 79 43 75
Response to Treatment	Responders None	10 23
5-year disease free survival	Yes No Censored (remaining are in Stage 4)	191 44 1 (33)

all cells positive within nucleus; and Nuclear Total Score defines the product of Nuclear % Intensity and nuclear % Positive.

Anti-EGFR antibody, anti-Her2/neu antibody, anti-phospho-EGFR antibody (Tyr845), and anti-phosphoHer2/neu

(Tyr877) were from Cell Signaling Technology, Inc. (Beverly, MA).

Study Cohort

Information of a breast cancer cohort was extracted from the Cell Signaling Database of Clinomics Biosciences Inc. The study cohort contained 269 breast tumor samples obtained from surgery. The patients had an average age of 63.7 years (ranging from 21 to 95 years). There were 24.9% with Stage 1, 45.0% with Stage 2, 17.8% with Stage 3, and 12.3% with Stage 4 breast cancer. Ninety-one patients (33.8%) were ER negative and 178 (66.2%) were ER positive. Forty-three (16.0%) patients were Her2/neu positive and 226 (84.0%) were Her2/neu negative (assayed by immunohistochemistry). There were 56.9% of patients received lumpectomy and node dissection, and 43.1% of patients received mastectomy and node dissection. Two hundred and two (75.1%) patients accepted CMF chemotherapy. One hundred and fifty-three (56.9%) patients went through localized (breast) radiation therapy. Among 33 patients who developed metastases, 10 patients were responders to the treatments, while 23 were none responders. Treatment response was defined according to RECIS [15]. Responders include complete response (CR) and partial response (PR), and none responders include stable disease (SD) and progressive disease (PR). Among the patients with Stage 0 to 3, 191 patients survived a 5-year disease-free interval, while 44 had recurrence within five years. The remaining patients' survival information was censored (details in Table **1**).

Clinical Outcomes

There were two outcomes included in the current analyses: disease free survival after therapy and treatment response. The experiments were designed to construct prognostic models to predict disease-free survival and treatment response in individual patients with breast cancers. The patient cohort was separated into two groups: one for patients with Stage 0 to 3 (Group I, $N = 235$), the other for patients with metastatic diseases (Group II, $N = 33$). All patients in Group II received CMF chemotherapy. For Group I, biomarkers were identified to build prognostic models for 5-year disease-free survival. For Group II, treatment response prediction models were constructed using a separately identified biomarker set.

Based on the clinical information in the studied cohort, we performed three analyses for both survival prediction and treatment response prediction, respectively (The details are listed in Results section). The clinical-pathological parameters for outcome predictions were chosen using gain Ratio Attribution Evaluation method with software WEKA. The parameters with merit greater than zero in the evaluation were included in the analysis.

Machine Learning Prediction System

Different machine learning algorithms perform differently according to datasets and the given tasks. The evaluation of different pathologic factors based on a single classifier may not lead to an objective conclusion in general. Therefore, we developed a prediction system combining three selected classifiers, a nearest neighbor method (*KStar*)

[16], Random Forests [17], and Neural Networks (*MultilayerPerceptron*) from the WEKA3.4 software package[1] [18]. These algorithms are accurate and stable in clinical outcome predictions on the studied cohort. The prediction accuracy of each classifier is evaluated by 10-fold cross validation, which has the least variance and bias of accuracy estimation among all validation methods including the leave-one-out method [19]. The average prediction accuracy of these classifiers is used as the evaluation criterion for breast cancer outcome predictions. If the prediction accuracy is unbalanced for a certain class label, the *F*-measure [20] was used as the evaluation criterion instead, which is defined as:

$F = 2 \times$ (recall $\times$ precision) / (recall + precision)

The average of the *F*-measure of all class labels was used as the final *F*-measure for a classifier. The average final *F*-measure of the three classifiers was used as the evaluation criterion of clinical outcome prediction (see Supplementary Information for details).

Statistical Methods

To assess the significance of the performance of an individual classifier, we computed the probability of the observed prediction accuracy occurring by chance (random prediction using a fair coin flip). The probability of doing at least as well as our prediction models by chance was calculated using *Binomial Distribution* functions [21] in software package *R* (http://www.r-project.org/). Statistical significance test was used to assess the significance of different prediction results on the studied cohort.

RESULTS

Identifying Specific High Risk Breast Cancer Patients

In order to identify specific high risk breast cancer patients, we used the clinical information in the studied cohort to predict whether or not a patient would survive a 5-year disease free interval. As we discussed previously, the traditional pathologic factors are inclined to be subjective and ER, PR, Her2/neu, as well as the demographic factors tend to be objective. The clinical information was separated into two categories: one includes the traditional pathologic factors and the other includes ER, PR, Her2/neu, plus the demographic factors. To compare the predictive power of these two different kinds of clinical information in a survival prediction, we conducted following analyses on the studied cohort. (1) Using pathologic factors (*histology, stage, pT, pN, and nodes positive*) to predict breast cancer survival. (2) Using *ER, PR, Her2/neu, age,* and *smoking* to predict breast cancer survival. (3) Using a comprehensive profile (*age, histology, stage, pT, pN, nodes positive, ER, PR, Her2/neu, and smoking*) to predict breast cancer survival. Our analyses demonstrated that the traditional pathologic factors entailed overall accuracy 94.6% in the survival prediction (Table **2**, Analysis 1) and the parameters of ER, PR, Her2/neu, plus patient's age and smoking status resulted in survival prediction accuracy 82% (Table **2**, Analysis 2), indicating that the traditional pathologic parameters are more accurate in the individualized risk assessment ($p < 0.000007$). However, when two kinds of predictive factors were incorporated, the survival prediction accuracy was further increased from 82%

---

[1] http://www.cs.waikato.ac.nz/ml/weka/

to 95.8% ($p < 0.000001$) for individual breast cancer patients (Table **2**, Analysis 3). It should also be pointed out that the addition of ER, PR, Her2/neu, plus patient's age and smoking status to the traditional pathologic factors did not significantly ($p > 0.05$) improve the prediction of the traditional factors alone. These results are consistent with the observations in Paik *et al.* [22]. Our results demonstrated that the identified comprehensive classifier can accurately define a patient's risk and be used to develop an optimal individualized therapy of breast cancer (see Supplement Information for detailed results).

Predicting Treatment Response of Each Individual Breast Cancer Patient

In order to assess a breast cancer patient's treatment response, we used the clinical information in the studied cohort to predict whether or not a patient is a responder to chemotherapy. It enables clinicians to determine the most suitable treatment options for each individual patient. To find out the best predictive parameters for the breast cancer treatment response, we conducted the following analyses on the studied cohort of the 33 patients with metastatic disease: (1) Using traditional pathologic factors (*histology, stage, pT, pN, nodes positive, metastasis site, surgery procedure,* and *chemotherapy*) to predict a patient's response to treatments. (2) Using a patient's ER, PR, Her2/neu, and demographic information (*ER, PR, Her2/neu, age, smoking, surgery procedure,* and *chemotherapy*) to predict a patient's response to treatments. (3) Using a comprehensive profile (*age, histology, stage, pT, pN, nodes positive, metastasis site, ER, PR, Her2/neu, smoking, surgery procedure,* and *chemotherapy*) to predict a patient's response to treatments. Using the traditional pathologic parameters, the accuracy of treatment response prediction is 94.7% (Table **3**, Analysis 1). Using ER,

PR, Her2/neu, and the patient's age and smoking status, the accuracy of treatment response prediction is 87.8% (Table **3**, Analysis 2). Again, the traditional pathologic factors entailed higher accuracy in the treatment response prediction ($p < 0.16$). However, when the parameters in Analysis 1 and Analysis 2 were combined, the treatment response prediction accuracy was increased from 87.8% to 95.3% ($p < 0.14$) (Table **3**, Analysis 3). Here, the addition of the biomarker panel to the traditional histological features did not significantly improve the prediction accuracy ($p > 0.05$). Our data demonstrate that the comprehensive profile encompassing both the traditional pathologic factors and ER, PR, Heu2/neu plus patient's age and smoking status was more accurate in predicting the treatment response of breast cancer patients. Our constructed classifier can accurately predict a patient's response to treatments, which enables the selection of the most suitable combination of therapeutic regimens for each individual breast cancer patient. Our results confirmed previous findings. Kendal previously reported that positive lymph node status is an independent risk factor for breast cancer prognosis [23]. Terry *et al.* identified the association between prolonged cigarette smoking and increased breast cancer risk [24].

Evaluating the Predictive Power of Total Her2/neu, Total EGFR, Phospho-Her2/neu, and Phospho-EGFR in Beast Cancer Survival

To more accurately reflect the functionalities of these two proteins *in vivo*. We performed the following analyses on the studied cohort to investigate the relative strength of the predictive power of total Her2/neu, total EGFR, phospho-EGFR and phospho-Her2/neu in breast cancer survival prediction. The six antibody measurements for each phosphorylated protein were used in disease-free survival predic-

Table 2. Breast cancer disease free survival prediction. Various breast cancer factors were analyzed for the predictive power in survival based on this study cohort (*N*=235). In each analysis, the clinical-pathological parameters used as predictors in the survival prediction were marked in the corresponding column. The average prediction accuracy of three methods, (*KStar*), Random Forests, and Neural Networks, was reported as the final result. *P* value represents the probability that the random prediction can perform at least as well as our prediction models

Survival predictors	Analysis 1	Analysis 2	Analysis 3
Age	x	x	x
Histology	x		x
pT	x		x
pN	x		x
Nodes Positive	x		x
Stage	x		x
ER		x	x
PR		x	x
Her2/neu		x	x
Smoking		x	x
Prediction accuracy	94.6%	82.0%	95.8%
Significance	$p < 1.3$ E-53	$p < 1.1$ E-25	$p < 7.2$ E-59

Table 3. **Treatment response prediction. Various breast cancer factors were analyzed for the predictive power in treatment response based on this study cohort (*N*=33). In each analysis, the clinical-pathological parameters used as predictors in the treatment response prediction were marked in the corresponding column. The average prediction accuracy of three methods, (*KStar*), Random Forests, and Neural Networks, was reported as the final result. *P* value represents the probability that the random prediction can perform at least as well as our prediction models**

Response predictors	Analysis 1	Analysis 2	Analysis 3
Age	x	x	x
Histology	x		x
Stage	x		x
pT	x		x
pN	x		x
Nodes Positive	x		x
Metastasis Site	x		x
ER		x	x
PR		x	x
Her2/neu		x	x
Smoking		x	x
Surgery procedure	x	x	x
Chemotherapy	x	x	x
Prediction accuracy	94.7%	87.8%	95.3%
Significance	$p < 3.9$ E-9	$p < 4.7$ E-7	$p < 4.0$ E-9

tion. As a control, we first used a patient's age, ER and PR status to predict the patient survival after therapy (Table **4**, Analysis 1). In this analysis, all patients were classified as low risk (5-year survival). It indicates that the predictive power of a patient's age, ER, and PR is weak, because they cannot identify any high risk patients. However, since majority of the patients (82.1%) survived 5 years in the studied cohort, the overall prediction accuracy is 82.1%. In this situation, we cannot use overall prediction accuracy as the evaluation criterion. As a resolution, we used the average *F*-measure [20] of the 5-year survival and non-5-year survival prediction results as the evaluation criterion. The *F*-measure is a combined measurement of *recall* and *precision*. It thus overcomes the problems caused by unbalanced prediction error rates. We then evaluated the survival prediction model by adding the expression level of total Her2/neu and total EGFR (Table **4**, Analysis 2), as well as phospho-Her2/neu

and phospho-EGFR (Table **4**, Analysis 3), respectively, to the prediction model. Our results showed that the survival prediction accuracy was increased with the addition of total Her2/neu and total EGFR measurements ($p < 0.15$). Furthermore, our analyses showed that the incorporation of phospho-Her2/neu and phospho-EGFR entailed the highest survival prediction accuracy ($p < 0.05$), while the incorporation of total Her2/neu, total EGFR, phospho-Her2/neu and phospho-EGFR resulted in the second highest survival prediction accuracy compared to the control ($p < 0.09$) (Table **4**, Analysis 4).

Association of Phospho-Heu2/neu and Phospho-EGFR in Survival Prediction of Individual Breast Cancer Patients

Both Her2/neu and EGFR belong to the EGFR family. There is growing evidence showing that the interaction

Table 4. **The predictive power of EGFR, Her2/neu, phospho-EGFR, phospho-Her2/neu in breast cancer survival assessment on this study cohort (*N*=235). Six parameters for antibody measurements were included in the analysis of phospho-EGFR, phospho-Her2/neu, and EGFR. *P* value represents the statistical significance of the improvement over the control analysis**

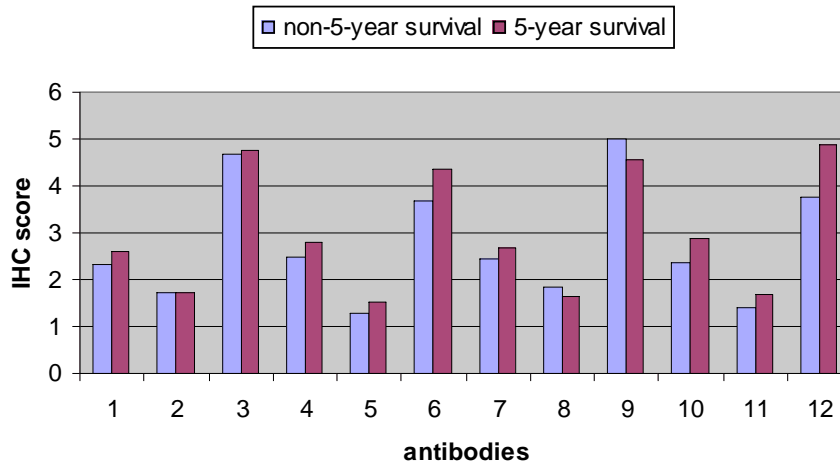| Analysis	Age	ER	PR	Her2/neu	pHer2/neu	EGFR	pEGFR	*F*-measure	Improvement
1	x	x	x					0.514	
2	x	x	x	x		x		0.561	$p < 0.15$
3	x	x	x		x		x	0.590	$p < 0.05$
4	x	x	x	x	x	x	x	0.575	$p < 0.09$

Fig. (1). Protein expression patterns of pEGFR and pHer2/neu in clinical samples. Antibody measurements: 1. pEGFR Cytoplasmic % Positive, 2. pEGFR Cytoplasmic % Intensity, 3. pEGFR Cytoplasmic Total Score, 4. pEGFR Nuclear % Positive, 5. pEGFR Nuclear % Intensity, 6. pEGFR Nuclear Total Score, 7. pHer2/neu Cytoplasmic % Positive, 8. pHer2/neu Cytoplasmic % Intensity, 9. pHer2/neu Cytoplasmic Total Score, 10. pHer2/neu Nuclear % Positive, 11. pHer2/neu Nuclear % Intensity, 12. pHer2/neu Nuclear Total Score.

among EGFR family members contributes to oncogenic transformation in human tumors. Studies found that human tumors are prone to co-express multi-EGFR family members, which are correlated to the more aggressive phenotypes, enhanced transforming activities, and a poor clinical outcome [25]. Therefore, we compared the predictive power of the combination of phospho-Her2/neu and phospho-EGFR vs. either phospho-Her2/neu or phospho-EGFR. The six antibody measurements for each phosphorylated protein were used in disease-free survival prediction. The protein expression patterns of phospho-EGFR and phospho-Her2/neu in clinical samples are shown in Fig. (**1**). Among the measured antibody scores, phospho-EGFR Cytoplasmic % Positive, phospho-EGFR Nuclear % Positive, phospho-EGFR Nuclear % Intensity, phospho-Her2/neu Cytoplasmic % Positive, phospho-Her2/neu Cytoplasmic % Intensity, phospho-Her2/neu Nuclear % Positive, and phospho-Her2/neu Nuclear % Intensity were differentially expressed among non-5-year survival and 5-year survival ($p < 0.05$). Furthermore, we found that age-ER-PR-phospho-Her2/neu-phospho-EGFR was more accurate in predicting the survival of breast cancer patients than either age-ER-PR-phospho-Her2/neu or age-ER-PR-phospho-EGFR (Table **5**). Our results imply that some interaction may exist between phospholated Her2/neu and phospholated EGFR in breast cancer. These observations are consistent with the clinical results in Her2/neu-targeted trastuzumab treatment for breast cancer patients. It was

found that some breast cancer patients with Her2/neu overexpression had no response to the treatment. *In vitro* experiments showed that the response to trastuzumab treatment was dependent not only on Her2/neu expression, but on the expression of other EGFR family member as well [26]. Our results are consistent with the recent publication showing that there is a simultaneous phosphorylation of Her2/neu and EGFR in Her2/neu overexpressed metastatic breast cancer [27]. Taken together, our results demonstrated that the combination of phospho-Her2/neu and phospho-EGFR is more accurate in survival prediction of individual breast cancer patients than either phospho-Her2/neu ($p < 0.10$) or phospho-EGFR ($p < 0.21$).

DISCUSSION

Breast cancer is a complex and heterogeneous disease, encompassing a wide range of pathologic entities and molecular profiles [28]. It is crucial for the physicians to accurately define a patient's risk of developing metastatic and recurrent diseases at diagnosis, which will determine the clinical course for the given patient, i.e. which patient should receive expensive and toxic adjuvant therapy, and which patient should avoid over-treatment. Furthermore, it is critical for the physicians to determine which combination of treatment is most suitable for each individual patient. However, it still remains challenging to make accurate predictive assessment of a patient's risk or response to certain treatment

Table 5. Comparison of the predictive power among phospho-Her2/neu, phospho-EGFR, and phospho-Her2/neu/phospho-EGFR in breast cancer survival assessment on this study cohort (*N*=235). *P* value represents the statistical significance of the improvement over the control analysis

Analysis	Age	ER	PR	Her2/neu	pHer2/neu	EGFR	pEGFR	*F*-measure	Improvement
1	x	x	x					0.514	
2	x	x	x		x			0.554	*p* < 0.19
3	x	x	x				x	0.533	*p* < 0.34
4	x	x	x		x		x	0.590	*p* < 0.05

regimens. This paper presents an accurate and robust model system for individualized survival and treatment response predictions for patients with breast cancer. Our results demonstrated that our model system can achieve overall accuracy above 95% in individualized breast cancer outcome predictions.

Previous studies of constructing predictive assessment models of breast cancer outcome were based on different criteria of patient selection and different data analysis methods. As different data analysis methods may generate different results leading to discrepant conclusions, it is difficult to perform generic comparisons of various predictive factors and ensure the repeatable accuracy of constructed outcome prediction models, making it difficult for the comprehensive evaluation and comparison of various predictive factors [28]. To develop a standardized and comprehensive profile for accurate clinical outcome predictions, we evaluated various factors in prediction assessments on the same patient cohort using a model system which combines three well known machine learning algorithms, *Kstar*, *random forests*, and *neural networks*. These three methods are all based on sound computational algorithms and have been widely deployed in various applications. They all provided accurate and stable results in this study. In order to provide an objective evaluation of various predictive factors in general, we used the average performance of these three methods as the evaluation criterion of the predictive assessments on the same patient cohort. It thus provides consistent, generic, and applicable evaluation with regard to the selection of patients and methods of data analysis. In addition, it ensures the accuracy and stability of the performance of the proposed prediction system despite the variability in data.

In this study, we also sought to apply new molecular knowledge and techniques in the evaluation of breast cancer predictive factors. Both EGFR and Her2/neu belong to the epidermal growth factor receptor (EGFR) superfamily. The phosphorylation in tyrosines is an essential step for the activation of EGFR and Her2/neu proteins, and is a strict requirement for these proteins to transform cells. Overexpression of EGFR and Her2/neu does not necessary indicate they are functionally activated. The binding of downstream proteins, such as Src, PI3K, Ras, Shc, and PLC-γ, to the tyrosine-phosphorylated EGFR and Her2/neu, induces a cascade of signal transduction events, ultimately leading to cell transformation [29]. Therefore, the expression levels of phosphorylation of EGFR and Her2/neu may more accurately reflect the status of these enzymes in the tumorgenesis of individual breast cancer patients. Indeed, studies of the correlation of EGFR overexpression with breast cancer prognosis and prediction have yielded conflicting results [8]. Some reports showed that the overexpression of EGFR was correlative to the adverse prognosis of breast cancer while other reports showed EGFR overexpression was not a good prognosis factor [8]. Study of the profiles of Her2/neu overexpression in a large group of patients found that patients with overexpression of unphosphoryled Her2/neu have a similar prognosis as those patients without the overexpression of Her2/neu [30]. These results suggest that the overexpression of phosphorylated EGFR and Her2/neu may be more relevant in predicting the clinical outcomes of individual breast cancer patients than the overexpression of total EGFR and Her2/neu. Our results showed that age-ER-PR-phospho-

Her2/neu-phospho-EGFR is more accurate in predicting the survival of individual breast cancer patients than age-ER-PR-Her2/neu-EGFR. Our results are consistent with the recently published observation, showing that both phosphorylated Her2/neu and phosphorylated EGFR are more efficient in predicting the clinical outcome of patients undergoing trastuzumab treatment [31].

In conclusion, we found that the incorporation of ER, PR, and Her2/neu, as well as other objective parameters, into the traditional pathologic factors is highly accurate to perform individualized survival and treatment response predictions for breast cancer patients. We also demonstrated that the measurements of active EGFR and Her2/neu *via* anti-phospho-antibodies are more predictive in the survival prediction than the measurements of overexpression of total EGFR and Her2/neu proteins. Nevertheless, it should be noted that the protein expression values were measured using immunohistochemistry assays, which are based on assessments of pathologists and are semi-quantitative. With the advances in protein array technology, the methodology presented in this study could be used to analyze the data generated using new technologies. Furthermore, it should be noted that the results presented in this paper need to be validated using independent datasets.

ACKNOWLEDGEMENTS

DISCLAIMER

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

[1]   Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Horto-bagyi GN. Is breast cancer survival improving? Cancer 2004; 100: 44-52.

[2]   Peto R, Boreham J, Clarke M, Davies C, Beral V. UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years. Lancet 2000; 355: 1822.

[3]   Baker SG, Kramer BS, Srivastava S. Markers for early detection of cancer: statistical guidelines for nested case-control studies. BMC Med Res Methodol 2002; 2: 4.

[4]   Schnitt SJ. Traditional and newer pathologic factors. J Natl Cancer Inst Monogr 2001; 22-26.

[5]   Hayes DF, Isaacs C, Stearns V. Prognostic factors in breast cancer: current and new predictors of metastasis. J Mamm Gland Biol Neoplasia 2001; 6: 375-392.

[6]   Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer 2005; 5: 845-856.

[7]   Olivotto IA, Bajdik CD, Ravdin PM, *et al*. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. J Clin Oncol 2005; 23: 2716-2725.

[8]   Ross JS, Linette GP, Stec J, *et al*. Breast cancer biomarkers and molecular medicine. Expert Rev Mol Diagn 2003; 3: 573-585.

[9]   Ross JS, Fletcher JA, Bloom KJ, *et al*. Targeted therapy in breast cancer: the HER-2/neu gene and protein. Mol Cell Proteomics 2004; 3: 379-398.

[10]   Warford A. Tissue microarrays: fast-tracking protein expression at the cellular level. Expert Rev Proteomics 2004; 1: 283-292.

[11] Warford A, Howat W, McCafferty J. Expression profiling by high-throughput immunohistochemistry. J Immunol Methods 2004; 290: 81-92.

[12] Kallakury BV, Sheehan CE, Ambros RA, *et al*. Correlation of p34cdc2 cyclin-dependent kinase overexpression, CD44s down-regulation, and HER-2/neu oncogene amplification with recurrence in prostatic adenocarcinomas. J Clin Oncol 1998; 16: 1302-1309.

[13] Ouban A, Muraca P, Yeatman T, Coppola D. Expression and distribution of insulin-like growth factor-1 receptor in human carcinomas. Hum Pathol 2003; 34: 803-808.

[14] Allred DC, Clark GM, Elledge R, *et al*. Association of p53 protein expression with tumor cell proliferation rate and clinical outcome in node-negative breast cancer. J Natl Cancer Inst 1993; 85: 200-206.

[15] Therasse P, Arbuck SG, Eisenhauer EA, *et al*. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst 2000; 92: 205-216.

[16] Dasarathy BV. Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques, 1991.

[17] Breiman L. Random Forests. Machine Learning 2001; 45: 5-32 .

[18] Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques (2nd Edition). Morgan Kaufmann, 2005.

[19] Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Proceedings of International Joint Conference on Artificial Intelligence (IJCAI) 1995; 1137-1143.

[20] Van Rijsbergen CJ. *Information Retrieval*. London: Butterworths, 1979.

[21] Staunton JE, Slonim DK, Coller HA, *et al*. Chemosensitivity prediction by transcriptional profiling. Proc Natl Acad Sci USA 2001; 98: 10787-10792.

[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 2004; 351: 2817-2826.

[23] Kendal WS. Lymph node-based prognostics: limitations with individualized cancer treatment. Am J Clin Oncol 2006; 29: 298-304.

[24] Terry PD, Miller AB, Rohan TE. Cigarette smoking and breast cancer risk: a long latency period? Int J Cancer 2002; 100: 723-728.

[25] Zaczek A, Brandt B, Bielawski KP. The diverse signaling network of EGFR, HER2, HER3 and HER4 tyrosine kinase receptors and the consequences for therapeutic approaches. Histol Histopathol 2005; 20: 1005-1015.

[26] Neve RM, Lane HA, Hynes NE. The role of overexpressed HER2 in transformation. Ann Oncol 2001; 12 (Suppl 1): S9-13.

[27] Gschwantler-Kaulich D, Hudelist G, Koestler WJ, *et al*. EGFR activity in HER-2 over-expressing metastatic breast cancer: evidence for simultaneous phosphorylation of Her-2/neu and EGFR. Oncol Rep 2005; 14: 305-311.

[28] Murphy N, Millar E, Lee CS. Gene expression profiling in breast cancer: towards individualising patient management. Pathology 2005; 37: 271-277.

[29] Holbro T, Hynes NE. ErbB receptors: directing key signaling networks throughout life. Annu Rev Pharmacol Toxicol 2004; 44: 195-217.

[30] Thor AD, Liu S, Edgerton S, *et al*. Activation (tyrosine phosphorylation) of ErbB-2 (HER-2/neu): a study of incidence and correlation with outcome in breast cancer. J Clin Oncol 2000; 18: 3230-3239.

[31] Hudelist G, Kostler WJ, Czerwenka K, *et al*. Her-2/neu and EGFR tyrosine kinase activation predict the efficacy of trastuzumab-based therapy in patients with metastatic breast cancer. Int J Cancer 2006, 118(5): 1126-34.

[32] American Joint Committee on Cancer Staging Manual. Philadelphia, PA: Lippincott-Raven, 1997.

SUPPLEMENTARY INFORMATION

"A Comprehensive Profile for Individualized Survival and Treatment Response Predictions in Patients with Breast Cancer: Involvements of Phospho-EGFR and Phospho-Her2/neu Proteins"

Lan Guo, Daniel C. Flynn, Vince Castranova, Xianglin Shi and Yong Qian

1. CLINICAL SAMPLE DESCRIPTION

The descriptions of the clinical samples ($N = 280$) used in this study are listed below. This cohort includes comprehensive pathological information, which contains traditional tumor markers (Diameter, Nodes Positive, pT, pN) and the more recent ones (ER, PR, and Her2/neu), as well as patients' demographic data.

Clinical-Pathological Parameters	Value	Occurrences
Tissue	Breast	234
	Adrenal	3
	Bone Marrow	1
	Brain	2
	Chest Wall	2
	Gastric	1
	Heart	2
	Kidney	2
	Liver	7
	Lung	1
	Lymph Node	19
	Spleen	2
	Testicle	2
	Thyroid	2
Histology	Infiltrating Ductal Carcinoma	173
	Lobular in situ Carcinoma	1
	Lobular Carcinoma	10
	Matched Benign Specimen	19
	Matched Benign Lymph Node	1
	Medullary Carcinoma	4
	Metastatic Breast Adenocarcinoma	22
	Metastatic Breast Carcinoma	1
	No Pathological Diagnosis	35
	Papillary Carcinoma	4
	Scirrhous Carcinoma	7
	Tubular Carcinoma	3
Surgery Procedure	Autopsy	27
	Biopsy	140
	Excision	2
	Mastectomy	107
	Splenectomy	2
	Thyroidectomy	2
Age	21 – 95	258
Sex	F	274
	M	6

Stage	0	9
	1	54
	2	98
	3	39
	4	3
Medical History	None	236
Chemotherapy	5FU	1
	CMF	68
	Cytotoxin	8
	Methotrexate	23
Radiation	Localized	236
ER	Negative	177
	Positive	59
PR	Negative	72
	Positive	164
Her2/neu	Negative	236
Metastasis Site	Node Abdominal	33
	None	203
Smoking	No	155
	Yes (> 20 packs/year)	40
	Yes (2 packs/day)	8
	Yes (3 packs/day)	33
Drug Dependency	None	236
pT	1	58
	2	89
	3	14
	4	32
pN	0	28
	1	48
	2	3
Nodes All	4 – 52	71
Nodes Positive	0 – 52	71
Diameter	1.0 cm – 10.0 cm	193
Resected Margin	Negative	34
	Positive	22
Response to Treatment	Responders	71
	None	199
5-year survival	Yes	202
	No	44

2. MACHINE LEARNING METHODS

In this study, we used three machine learning classifiers, *Kstar, RandomForest*, and *MultilayerPerceptron* implemented in WEKA [1] (http://www.cs.waikato.ac.nz/ml/weka/), to build the clinical outcome prediction system for patients with breast cancer. Specifically, *Kstar* is a nearest neighbor method with a generalized distance function based on transformations weighted by probabilities. A random forest is an ensemble of hundreds or thousands of classification trees built on bootstrapped samples.

The final classification result is obtained by majority voting of these decision trees. This method is considered to be more robust than a single classification tree. A *MultilayerPerceptron* is a neural network that trains the prediction model using back-propagation, which is a process of adjusting weights in the network connections to decrease the number of errors. These three methods are based on different computational algorithms and have been widely employed in various applications. They all provided accurate and stable results in this study. In order to provide an objective and general evaluation of various prediction factors, we used the average performance of these three methods as the evaluation criterion of the predictive assessments. The performance of each method is evaluated using 10-fold cross validation. In the 10-fold cross validation, the data set was randomly partitioned into 10 folds of equal size with possible exception of the last fold (the last fold contains the remaining samples). The prediction models were trained and tested 10 times. Each time, 9 folds were picked to build the prediction model, while the remaining fold was validated on the prediction model. We used 10-fold cross validation to evaluate the prediction models in this study, because the estimation accuracy by this validation method has been proven to have the lowest bias and variance among all validation methods, including the leave-one-out method [2]. It thus provides an objective evaluation of the performance of our prediction models in general.

3. PREDICTION ACCURACY MEASUREMENTS

Two prediction accuracy measurements were used in this study, overall prediction accuracy and the F-measure. The overall prediction accuracy is the number of patient samples whose specific clinical outcome was correctly predicted by the proposed model divided by the total number of patient samples. In our study, we also encountered unbalanced prediction results. For instance, when we used *age, ER*, and *PR* to predict patient survival, all high risk patients (none 5-year survival) were misclassified as low risk (5-year survival). Although the overall prediction accuracy is high (82.1%), such prediction models are not useful for clinical applications. In the situations where the prediction accuracy is unbalanced for a certain class label (as we just described), the *F*-measure [3] was used as the evaluation criterion. For a certain class label, the *F*-measure is defined as:

$$F = 2 \times (\text{recall} \times \text{precision}) / (\text{recall} + \text{precision})$$

where *recall* is the ratio of the number of correct cases to the total number of the cases in the actual cohort, and *precision* is the ratio of the number of correct cases to the total number of correct cases and incorrect cases in the prediction results. The average of the *F*-measure of all (in our case two) class labels was used the final *F*-measure for a classifier.

4. DETAILED PREDICTION RESULTS IN THE STUDY

I. Individualized Survival Prediction Using Various Prediction Factors

a. Predictors of patient 5-year survival

Histology, pT, pN, Node Positive, Resected Margin, Tumor Diameter, and Stage.

Classifier	5-year Survival Prediction Accuracy (*N*=202)	None 5-year Survival Prediction Accuracy (*N*=44)	Overall Prediction Accuracy (*N*=246)
Kstar	98.5%	84.1%	95.9%
RandomForest	99.0%	72.7%	94.3%
Neural Net	97.5%	75.0%	93.5%
Average	98.3%	77.3%	94.6%

b. Predictors of patient 5-year survival

Age, Medical History, Smoking, Drug Dependency, ER, PR, and Her2/neu status

Classifier	5-year Survival Prediction Accuracy (*N*=202)	None 5-year Survival Prediction Accuracy (*N*=44)	Overall Prediction Accuracy (*N*=246)
Kstar	99.5%	9.1%	83.3%
RandomForest	91.6%	36.4%	81.7%
Neural Net	95.5%	18.2%	80.9%
Average	95.5%	21.2%	82.0%

c. Predictors of patient 5-year survival

Age, Histology, pT, pN, Node Positive, Resected Margin, Tumor Diameter, Stage, Medical History, Smoking, Drug Dependency, ER, PR, and Her2/neu status.

Classifier	5-year Survival Prediction Accuracy (N=202)	None 5-year Survival Prediction Accuracy (N=44)	Overall Prediction Accuracy (N=246)
Kstar	97.0%	93.2%	96.3%
RandomForest	98.5%	81.8%	95.9%
Neural Net	97.0%	88.6%	95.5%
Average	97.5%	87.9%	95.8%

II. Individualized Treatment Response Prediction Using Various Prediction Factors

a. Predictors of patient response to treatments

Histology, tissue, pT, pN, Node All, Node Positive, Resected Margin, Tumor Diameter, Stage, Metastasis Site, Surgery Procedure, Chemotherapy, and Radiation.

Classifier	Responder Prediction Accuracy (N=71)	Non-Responder Prediction Accuracy (N=199)	Overall Prediction Accuracy (N=280)
Kstar	83.1%	99.5%	95.2%
RandomForest	80.3%	99.0%	94.1%
Neural Net	88.7%	97.0%	94.8%
Average	84.0%	98.5%	94.7%

b. Predictors of patient response to treatments

Age, Medical History, Smoking, Drug Dependency, ER, PR, Her/2neu status, Surgery Procedure, Chemotherapy, and Radiation.

Classifier	Responder Prediction Accuracy (N=71)	Non-Responder Prediction Accuracy (N=199)	Overall Prediction Accuracy (N=280)
Kstar	60.6%	97.5%	87.8%
RandomForest	62.0%	95.0%	86.3%
Neural Net	77.5%	93.5%	89.3%
Average	66.7%	95.3%	87.8%

c. Predictors of patient response to treatments

Age, Histology, tissue, pT, pN, Node All, Node Positive, Resected Margin, Tumor Diameter, Stage, Metastasis Site, Medical History, Smoking, Drug Dependency, ER, PR, Her2/neu status, Surgery Procedure, Chemotherapy, and Radiation.

Classifier	Responder Prediction Accuracy (N=71)	Non-Responder Prediction Accuracy (N=199)	Overall Prediction Accuracy (N=280)
Kstar	91.5%	99.0%	97.0%
RandomForest	87.3%	95.5%	93.3%
Neural Net	91.5%	97.0%	95.6%
Average	90.1%	97.2%	95.3%

REFERENCES

[1]   Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques (2nd Edition). 2005. Morgan Kaufmann.
[2]   Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. International Joint Conference on Artificial Intelligence (IJCAI). 1995;1137-1143.
[3]   Van Rijsbergen CJ. Information Retrieval. 1979. Butterworths, London.