Prediction of Nodal Involvement in Breast Cancer Based on Multiparametric Protein Analyses from Preoperative Core Needle Biopsies of the Primary Lesion

Georg Sauer,1 Nicole Schneiderhan-Marra,2 Cornelia Kazmaier,2 Kathrin Hutzel,2 Karin Koretz,3 Rainer Muche,4 Rolf Kreienberg,1 Thomas Joos,2 and Helmut Deissler1

Abstract

Purpose: Identification of molecular characteristics that are useful to define subgroups of patients fitting into differential treatment schemes is considered a most promising approach in cancer research. In this first study of such type, we therefore investigated the potential of multiplexed sandwich immunoassays to define protein expression profiles indicative of clinically relevant properties of malignant tumors.

Experimental Design: Lysates prepared from large core needle biopsies of 113 invasive breast carcinomas were analyzed with bead-based miniaturized sandwich immunoassays specific for 54 preselected proteins.

Results: Five protein concentrations [fibroblast growth factor-2 (FGF-2), Fas, Fas ligand, tissue inhibitor of metalloproteinase-1, and RANTES] were significantly different in the patients with or without axillary lymph node metastasis. All 15 protein parameters that resulted in P values <0.2 and other diagnostic information [estrogen receptor (ER) status, tumor size, and histologic grading] were analyzed together by multivariate logistic regression. This yielded sets of five (FGF-2, Fas, Fas ligand, IP10, and PDGF-AB/BB) or six (ER staining intensity, FGF-2, Fas ligand, matrix metalloproteinase-13, PDGF-AB/BB, and IP10) parameters for which receiver-operator characteristic analyses revealed high sensitivities and specificities [area under curve (AUC) = 0.75 and AUC = 0.83] to predict the nodal status. A similar analysis including all identified parameters of potential value (15 proteins, ER staining intensity, T) without selection resulted in a receiver-operator characteristic curve with an AUC of 0.87.

Conclusion: We clearly showed that this approach can be used to quantify numerous proteins from breast biopsies accurately in parallel and define sets of proteins whose combined analyses allow the prediction of nodal involvement with high specificity and sensitivity.

The clinical management of breast cancer has been substantially improved by the systematic use of optimized strategies of systemic adjuvant treatment. As a consequence, the chance of dying from breast cancer decreases despite an increasing incidence of this disease. For all types of local and systemic therapy, however, it is difficult to predict the individual patient’s immediate benefit and the effects of optional treatment modalities on long-term prognosis. This results in substantial number of patients who either miss the chance to receive highly efficient specific treatment or are subject to substantial overtreatment not resulting in any therapeutic benefit. Therefore, identification of characteristics of individual tumors that are useful to define subgroups of patients fitting into differential treatment schemes is considered a most promising approach in cancer research to improve the clinical outcome. To identify useful indicators of prognosis and therapeutic response, polymorphisms defining the genetic background of the patients, tumor-specific genomic alterations acquired during carcinogenesis, and microarray-derived comprehensive gene expression profiles were analyzed in numerous studies published in the last years. Their results dramatically increased the knowledge of interindividual molecular differences of breast tumors and provided the basis for an individualization of clinical care. It is evident and widely accepted now that it is possible to classify cases of breast cancer according to specific gene expression signatures that most likely reflect relevant features of the tumors better than cell and tissue morphology (1, 2). During the last years, methods for a parallelized quantification of proteins in cell or tissue extracts by means of multiplexed sandwich immunoassays have become available (3, 4) and their value as tools to define protein expression profiles indicative of

Authors’ Affiliations: 1Department of Obstetrics and Gynaecology, University of Ulm, Germany; 2Natural and Medical Sciences Institute at the University of Tübingen, Reutlingen, Germany; and 3Department of Pathology and 4Department of Biometry, University of Ulm, Ulm, Germany

Received 11/6/07; revised 1/17/08; accepted 2/8/08.

Grant support: German Ministry for Research and Education through German Federal Ministry of Education and Research grants 0312879A and 0312879D. The costs of the publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: G. Sauer and N. Schneiderhan-Marra contributed equally to this work.

Requests for reprints: Helmut Deissler, Department of Obstetrics and Gynaecology, University of Ulm Medical School, Frauensteige 14, D-89075 Ulm, Germany. Phone: 49-731-500-58660; Fax: 49-731-500-58804; E-mail: helmut.deissler@uniklinik-ulm.de.

©2008 American Association for Cancer Research. doi:10.1158/1078-0432.CCR-07-4802
clinically relevant properties of malignant tumors is currently evaluated (5). Although planar antibody arrays can be used in analogy to those used for mRNA profiling, bead-based methods with antibodies immobilized on the surface of color-coded microspheres are considered interesting alternatives, especially when a limited number of parameters have to be determined in parallel from many samples, because in terms of sensitivity, reliability, and accuracy, these assays are similar to well-established ELISAs done in standard microliter plates.

In this study, we analyzed lysates prepared from large core needle biopsy specimens of 113 invasive breast carcinomas with bead-based miniaturized sandwich immunoassays specific for 54 proteins of potential prognostic or predictive value that had been selected on the basis of published data. Because the histologic assessment of image-guided large core needle biopsies is a well-established reliable method to initially characterize breast tumors (6, 7) and provide the basis of decisions on subsequent therapy, additional information gained by molecular characterization should be most valuable at this stage. We show that this approach can be used to quantify numerous proteins from breast biopsies accurately in parallel and define sets of proteins whose combined analyses allow the prediction of nodal involvement with high specificity and sensitivity.

Materials and Methods

Patient characteristics. In addition to tissue samples used to characterize the tumors by conventional histopathologic assessment and immunohistochemical staining of routine parameters [Her-2, estrogen receptor (ER), and progesterone receptor], biopsies for multiparametric protein analyses were taken from 124 patients referred to the Breast Centre of the University of Ulm with breast lesions that had been initially detected by mammography and/or ultrasound imaging. Based on histologic criteria, 113 of these were classified as invasive mammary carcinomas. One carcinoma in situ, eight benign breast tumors, two tumors of other cellular origins (plasmacytoma, lymphoepithelioma), and two invasive carcinomas of which clinical data were incomplete were excluded from this study that was based on the remaining 111 biopsies of invasive breast carcinomas. The core variables of this group of patients are summarized in Table 1. The diagnostic routine parameters ER, progesterone receptor, and Her-2 (assessed by immunohistochemical staining); grading G; and size T of the tumors were available to be included in multivariate analyses.

Image-guided large core needle biopsies of breast lesions. We did a two-dimensional ultrasound examination (Voluson 730, Kretztechnik through GE Medical Systems; 3D-US transducer 5-13 MHz, 30° volume sector) with the patients in supine position and elevated arms to localize the primarily detected tumor or additional lesions in the same or the contralateral breast. Before the first breast biopsy, an informed consent was obtained (under approved internal review board protocol 130/2003; University of Ulm) and a history of blood coagulation problems was requested from each patient. No laboratory tests were done unless the patient was under anticoagulation therapy or reported a history of coagulopathy. All biopsies were taken by means of a core needle throw (22-mm excursion) automated biopsy gun (Bard-Magnum Biopsy Instrument) fitted with 14-gauge needles. After the needle had been placed at the edge of the lesion under 2D-US guidance (prefiring position), the 22-mm core needle throw was executed. Then, a 3D-US volume data set (10 mb) was acquired and converted into a multiplanar imaging display to define the precise postfiring position of the needle. All biopsies used in this study to analyze protein concentrations in tissue lysates resulted from central hits of the tumor mass. Taken biopsies were immediately frozen in liquid nitrogen and stored at -80°C.

Preparation of tissue lysates. Lysis buffer was prepared by supplementing 50 mmol/L Tris-Cl (pH 7.5), 400 mmol/L NaCl, 1 mmol/L CaCl2, 1 mmol/L MgCl2, 1% Triton X-100 with protease inhibitors, and phosphatase inhibitors I and II from Sigma. Of this buffer, 10 μL/mg of tissue was frozen and ground (1 min, 2,000 min⁻¹) together with the frozen biopsy in a microdisembranator (Braun Biotech). The resulting tissue-buffer powders were transferred to fresh prechilled reaction tubes and stored at -80°C. To prepare tissue extracts, these powders were thawed and, interrupted by occasional vortexing, kept on ice for 1 h. After centrifugation for 10 min at 15,000 × g, 4°C, the supernatant was separated from the cellular debris, snap frozen in suitable portions, and stored at -80°C. Protein concentrations were determined by the Lowry method (reagents from Bio-Rad) in relation to standard solutions containing bovine serum albumin.

Bead-based protein microarrays. To measure the concentrations of proteins in tissue lysates, we used three multiplex assays based on the “Beadlit Human Flex Kits” and established additional sandwich immunoassays with pairs of analyte-specific antibodies from Abcam, Bender MedSystems, Biosource, NeoMarkers/LabVision, R&D Systems, or Technoclone. For each assay, one specific antibody was coupled with color-encoded microspheres and used in combination with a second specific antibody that had been labeled with biotin to enable detection with phycoerythrin-streptavidin. Several pairs of antibodies were combined to multiplex assays for parallel analyses of different analytes (Fig. 1).

The carboxylated color-encoded microspheres (Luminex X-MAP) from Luminex were activated and coated with capture antibodies by a standard procedure including carboxidimide and succinimid reactions. After coupling of the antibody, beads were washed with 0.1% Tween 20 in PBS (pH 7.4) and resuspended at a maximum density (complete recovery after coupling) of 2.5 × 10⁹/mL in PBS containing 1% bovine serum albumin and 0.05% sodium azide. These suspensions were stored at 2°C to 8°C. Efficient coupling was confirmed by detection of bound antibodies with phycoerythrin-labeled antibodies recognizing their constant regions.

The detection antibodies specific for the proteins to be analyzed were either purchased as biotin conjugates or biotinylated with NHS-LC-LC-Biotin, of which a 50-fold molar excess was allowed to react in PBS for 2 h on ice. After purification by size exclusion chromatography with spin columns from Princeton Separations, the antibody solution was stored at the same volume glycerol and stored at -20°C.

The sandwich immunoassays were done in 96-well filter plates processed in a thermomixer. A total amount of 2 or 10 μg of protein solubilized from breast biopsies was diluted in 30 μL assay buffer (AB; blocking reagent for ELISA), mixed with 30 μL bead suspension, and incubated for 2 h at room temperature or overnight at 4°C to allow binding of the antigens. After each of the following steps, the solution was drained off by means of a Multiscreen Vacuum Manifold and the beads were washed twice. Subsequently, the mixture of biotinylated antibodies (2 h at room temperature) and R-phycoerythrin–labeled streptavidin were added. After the final washing step, beads were resuspended in 100 μL AB and analyzed in a Luminex 100 IS (Luminex). Median fluorescence intensities of at least 100 beads per assay were recorded for each of the samples that were measured in duplicates. Absolute quantification was based on an 8-point standard dilution series made with recombinant proteins purchased from Affinity BioReagents, Calbiochem/EMD Biosciences, R&D Systems, or Technoclone.

Data analyses and statistical models. The value of individual parameters as predictors of nodal involvement was initially analyzed with the Mann-Whitney test using Statistica 7.1 (StatSoft) and univariate logistic regression with Statistical Analysis System software (SAS Institute), which was also used for multivariate modeling and subsequent calculations.

Parameters that were independent, as revealed by multicollinearity analysis, and reached P values <0.2 in univariate logistic regression,
were used to calculate models based on multiple logistic regression that were shaped by sequential backward elimination of the least influential parameters in each step. Based on the resulting sets of parameters, receiver-operator characteristic analyses were done to show their sensitivities and specificities in the prediction of nodal status.

Results

Design and generation of an assay system to analyze protein expression in breast large core needle biopsies. The main purpose of this study was to prove the feasibility of miniaturized and multiplexed sandwich immunoassays in diagnostic applications based on breast large core needle biopsies or small amounts of tissue of other origin. We included proteins with reported or potential prognostic or predictive value in breast cancer for which suitable pairs of antibodies and recombinant antigens to generate standard curves were available. In addition, commercially available multiplexed assays to measure cytokines and growth factors taken could be included in subsequent processing and data analyses, indicating that biopsy-based protein quantification with multiplexed sandwich immunoassays is feasible and sufficiently robust.

Correlation of measured individual parameters with nodal status of the patients. All of the 36 informative parameters were analyzed for a potential value in the prediction of lymph node involvement in patients with histopathologically confirmed invasive breast carcinomas (n = 111), two patients were excluded because of unclear nodal status; Table 1) with the Mann-Whitney test and univariate logistic regression. To verify the normality of the group of patients in this study and to gain additional parameters for subsequent multivariate modeling, immunohistochemically assessed ER, PR, and Her-2; histologic grading (G); and tumor size (T) were included in this analysis. As shown in Table 2, both univariate logistic regression and the Mann-Whitney test revealed significant (P < 0.05) correlations of the measured concentrations of fibroblast growth factor-2 (FGF-2), Fas, and Fas ligand with nodal status, whereas the P values of TIMP-1 and RANTES were <0.05 only in one of these calculations. Box-whisker plots to visualize the differences

\[ \text{N. Schneiderhan-Marra, et al., submitted for publication.} \]
in the measured concentrations of TIMP-1, Fas, FGF-2, and RANTES between the groups of patients with (N1) and without (N0) involvement of axillary lymph nodes are shown in Fig. 2. In addition, tumor size (T) and grading (G) were confirmed to be significantly higher in patients with nodal involvement. By univariate logistic regression to analyze differences in the N0 and N1 groups, a total of 15 measured protein concentrations, T, G, and the intensity of immunohistologic staining of the ER resulted in P values <0.21 and were therefore used in subsequent multivariate analyses.

### Prediction of nodal involvement based on a set of parameters.

Multivariate logistic regression was used to evaluate the predictive value of the 15 immunoassay-determined protein concentrations for which univariate analyses had revealed a potential influence (P < 0.21). In parallel calculations, histologic tumor size T and intensity of immunohistochemical staining of the ER were added as additional factors. The resulting models based on these sets of 15 or 17 parameters allowed the prediction of nodal involvement with high sensitivity and specificity as indicated by area-under-curve (AUC) values of the receiver-operator characteristic diagrams of 0.80 [95% confidence interval (95% CI), 0.71-0.89] and 0.87 (95% CI, 0.80-0.94), respectively (Fig. 3). By a process of backward selection in which the least influential variables were removed in each step before recalculation of the model, small sets of five or six variables with only slightly reduced performance (AUC, 0.75; 95% CI, 0.66-0.85 and AUC, 0.83; 95% CI, 0.74-0.91) were defined (Fig. 3). To estimate the performance of other suboptimal combinations of parameters that might be chosen for nonstatistical reasons like robustness of the assays, various models were calculated with a limited number of input variables. For example, the combination of the four assays to measure FGF-2, matrix metalloproteinase (MMP)-13, PDGF-AB/BB, and TIMP-1 yielded a receiver-operator characteristic curve with an AUC of 0.72 (95% CI, 0.62-0.82). In view of potential diagnostic applications, the different assay systems were compared in terms of specificity at constant high sensitivities of 80% and 90% (Fig. 4). Corresponding sensitivities of about 50% indicate that even this initial feasibility study has yielded parameter sets that can be used to generate information highly relevant in clinical practice.

### Discussion

This study was initiated to prove the feasibility of novel diagnostic approaches with bead-based multiplexed sandwich immunoassays used to analyze protein expression in large core needle–derived breast tumor tissues. Remarkably, in addition to clearly achieving this goal, we were able to show that by applying this methodology, parameter sets of high clinical relevance can be defined and narrowed down to few factors that can easily be analyzed in validation studies with a high number of patients and as a diagnostic routine.

Compared with expression profiling based on mRNA levels, in which the whole transcriptome or every subgroup of transcripts considered relevant to the studied question can be covered by a robust and uniform technique, the selection of...
The expression of mRNA encoding FGF-2, also known as basic FGF, was recognized to be lower in malignant breast tumors than in benign lesions and normal tissue (8, 9). Low levels of the protein in breast cancer tissue were found to be associated with poor prognosis (10), which was confirmed in a large study also revealing a correlation with the absence of axillary metastasis (11). In accordance with these results, FGF-2 was the single factor with the highest value to predict spread to the lymph nodes in our group of patients harboring invasive carcinomas. However, it is not clear how down-regulation of FGF-2 might contribute to tumor progression because from its roles as a growth-stimulating factor, acting directly or by stimulation of shedding of syndecan-1 from stromal fibroblasts (12), or proangiogenic peptide, a direct rather than indirect correlation would have been expected. Recently, Korah et al. (13) have suggested that low levels of FGF-2 may cause loss of laminin 5 and, as a consequence, dedifferentiation of mammary ductal structures, but further experimental evidence is needed to support this concept. The observed differences in the expression of PDGF-AB/BB, Fas/Fasl, and MMP-13 (collagenase-3) were in accordance with their mechanisms of action and reported prognostic relevance (14–18). That the parameter sets that resulted from multivariate logistic regression included MMP-13 was surprising because of few biopsies of N0 patients in which strong MMP-13 expression led to a higher mean value in the N0 group and a quite low difference between the median values calculated to be higher in the N1 group. The chemokines RANTES (CCL5) and IP-10 (CXCL10), respectively, clearly contributed to the predictive values of parameter sets defined by multivariate analyses. Their potential roles in breast cancer were suggested on the basis of experiments with cell lines (19, 20). RANTES was described by Yaal-Hahohen et al. (21) as a factor predicting disease progression of stage II breast cancer patients. Interestingly, expression of this cytokine, in combination with the assessment of the ER status, provided improved strength in the prediction of disease progression. This is in

### Table 2. Results of univariate logistic regression and the Mann-Whitney test to identify parameters of value to predict nodal status

<table>
<thead>
<tr>
<th>Parameter</th>
<th>N0 group (ng/mL)</th>
<th>N1 group (ng/mL)</th>
<th>P (Mann-Whitney)</th>
<th>P (logistic regression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Median</td>
<td>Mean ± SD</td>
<td>Median</td>
</tr>
<tr>
<td>1 FGF-2</td>
<td>1.96 ± 1.90</td>
<td>1.42</td>
<td>1.08 ± 0.86</td>
<td>0.79</td>
</tr>
<tr>
<td>2 Fas</td>
<td>0.39 ± 0.25</td>
<td>0.31</td>
<td>0.28 ± 0.24</td>
<td>0.21</td>
</tr>
<tr>
<td>3 Fasl ligand</td>
<td>0/1-Encoded</td>
<td>0/1-Encoded</td>
<td>0.0236</td>
<td>0.0277</td>
</tr>
<tr>
<td>4 TIMP-1</td>
<td>33.65 ± 24.98</td>
<td>25.06</td>
<td>24.19 ± 17.71</td>
<td>17.65</td>
</tr>
<tr>
<td>5 MMP-13</td>
<td>1.81 ± 2.97</td>
<td>0.41</td>
<td>0.90 ± 1.15</td>
<td>0.48</td>
</tr>
<tr>
<td>6 MMP-1</td>
<td>28.44 ± 55.22</td>
<td>5.33</td>
<td>13.24 ± 18.69</td>
<td>6.15</td>
</tr>
<tr>
<td>7 TIMP-2</td>
<td>14.11 ± 10.36</td>
<td>11.89</td>
<td>11.45 ± 6.67</td>
<td>11.10</td>
</tr>
<tr>
<td>8 PDGF-AB/BB</td>
<td>0.23 ± 0.26</td>
<td>0.20</td>
<td>0.31 ± 0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>9 TNF-2</td>
<td>0.44 ± 0.44</td>
<td>0.31</td>
<td>0.34 ± 0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>10 IP-10</td>
<td>8.05 ± 14.4</td>
<td>2.81</td>
<td>4.66 ± 6.13</td>
<td>2.78</td>
</tr>
<tr>
<td>11 RANTES</td>
<td>1.19 ± 1.07</td>
<td>0.82</td>
<td>1.52 ± 1.00</td>
<td>1.07</td>
</tr>
<tr>
<td>12 MMP-10</td>
<td>0/1-Encoded</td>
<td>0.3015</td>
<td>0.1936</td>
<td></td>
</tr>
<tr>
<td>13 MIP1a</td>
<td>0/1-Encoded</td>
<td>0.1818</td>
<td>0.1997</td>
<td></td>
</tr>
<tr>
<td>14 MMP-7</td>
<td>4.93 ± 15.14</td>
<td>0.43</td>
<td>1.81 ± 3.51</td>
<td>0.47</td>
</tr>
<tr>
<td>15 ICAM-1</td>
<td>92.93 ± 42.18</td>
<td>93.00</td>
<td>87.71 ± 47.14</td>
<td>84.35</td>
</tr>
<tr>
<td>ER staining intensity</td>
<td>1.79 ± 1.15</td>
<td>2.0</td>
<td>1.82 ± 1.32</td>
<td>2.5</td>
</tr>
<tr>
<td>T (tumor size)</td>
<td>1.60 ± 0.75</td>
<td>1.5</td>
<td>2.03 ± 0.75</td>
<td>2.0</td>
</tr>
<tr>
<td>G (grading)</td>
<td>2.05 ± 0.53</td>
<td>2.0</td>
<td>2.38 ± 0.55</td>
<td>2.5</td>
</tr>
</tbody>
</table>

NOTE: Only factors with P values <0.21, which were used as input variables in subsequent multivariate logistic regression, are shown.
accordance with the results of multivariate modeling in this study, which showed a substantial contribution of the ER staining intensity in model calculations including such basic parameters.

The molecular differences of defined subclasses of breast cancers are strongly related to prognosis and response to therapeutic options. To address critical decisions in clinical practice, genome-wide expression profiling was used to distinguish defined (sub)groups of patients with different prognosis or response to specific treatment, resulting in sets of a limited number of informative parameters. This strategy yielded two, only to some extent overlapping, analytical systems to predict the occurrence of distant metastasis and the patients' long-term prognosis by quantifying 70 or 74, respectively, tumor-expressed mRNAs (22–24). Based on investigations of small groups of patients, gene expression signatures predictive of complete response to neoadjuvant chemotherapy (25) or recurrence after mastectomy (26, 27) were established, but their diagnostic values still have to be confirmed in studies with more patients. In addition, conventional molecular and histologic features of breast tumors were defined by characteristic expression patterns. The signature of expression of 97 genes that well correlated with histopathologically assessed grading allowed a subclassification of grade 2 tumors into two groups with high or low risk of progression (28). In all these studies, informative genes were defined by the results of comprehensive array-based profiling. Paik et al. (29) showed that rational preselection of potentially relevant genes is a promising strategy to find useful signatures without screening of the whole transcriptome: reverse transcription-PCR analyses of mRNAs of only 16 genes, extracted from paraffin-embedded tumor tissues, allowed the prediction of recurrence of node-negative, tamoxifen-treated breast cancer and a subclassification of patients into groups with different risks of relapse. Despite such interesting results obtained from paraffin-embedded material, the tumor tissues used for mRNA profiling usually were taken during surgery. In addition, fine-needle aspirations and core needle biopsies were confirmed as potential sources of tissue (30–32).

Investigating mRNA profiles is attractive because the physicochemical uniformity of all mRNA species expressed by a cell allows their analyses with standard methods of relatively low complexity that had to be refined to meet the requirements of the parallel assessment of a high number of analytes. Similar methodologic approaches are difficult to establish for the proteome due to the structural heterogeneity of proteins and their variable stability. On the other hand, normal and malignant phenotypes are primarily a result of protein functions and, therefore, their analyses might yield valuable information that cannot be extracted from RNA data. In a retrospective study similar to the reverse transcription-PCR
analyses by Paik et al. (29), immunohistochemical staining of 1,600 tumor tissues with antibodies specific for 26 selected proteins yielded a set of 21 proteins whose expression pattern highly significantly correlated with metastasis-free survival as the most important determinant of clinical outcome (33). An additional approach based on immunohistochemical staining with 163 antibodies yielded sets of 5 or 6 (dependent on the model) antigens whose assessment allowed to identify ER-positive, but not ER-negative, breast cancer patients with poor clinical outcome (34). These investigations proved that molecular characteristics of breast lesions, defined at the mRNA or protein level by screening of candidate factors or arrays covering high numbers of different analytes, can be used, in addition to conventional classification and staging, as indicators of prognosis and associated clinical attributes like nodal status and to predict response to the therapeutic options. In this first study of this type, we clearly showed that bead-based multiplexed sandwich immunoassays are suitable and robust enough to be used for screening of potentially relevant factors to yield parameter sets of clinical value.

Fig. 3. Summary of results of multivariate logistic regression to evaluate the value of sets of parameters to predict nodal involvement. Data included in the analysis by multivariate logistic regression were from 15 immunoassays [supplemented in parallel calculations with staging T and ER staining intensity (ER-SI)] for which low (<0.2) P values had been determined by univariate comparison of the N0 and N1 groups (see Table 2). Diagrams show the receiver-operator characteristics of these basic sets without parameter selection (top graphs) and small parameter sets derived by stepwise backward selection (bottom graphs). The high AUC values, indicative of the performance as a diagnostic test, of the initial sets [0.80 (95% CI, 0.71-0.89) and 0.87 (95% CI 0.80-0.94)] were essentially conserved in the smaller sets of variables [0.75 (95% CI, 0.66-0.85) and 0.83 (95% CI 0.74-0.91)]. The AUC of the best single predictive parameter FGF-2 was calculated to be 0.67 (95% CI, 0.57-0.77).
The data set generated in this study by parallel assessment of 54 protein concentrations for each of the 111 biopsies from invasive breast cancer was supplemented with diagnostic core parameters and analyzed by univariate and multivariate logistic regression to identify predictors of axillary lymph node metastasis. When the derived sets of few parameters are compared with the best single-parameter FGF-2 having a quite high predictive value calculated by univariate analysis, it is a striking result that even few combined parameters are much more powerful predictors of nodal involvement (Fig. 4). As axillary lymph node dissection is one of the main causes of surgical morbidity in breast cancer patients, a reliable exclusion from this treatment would be highly beneficial to these patients. Especially in patients with smaller tumors (T1, T2) of which <50% harbor positive lymph nodes (35), the number of prophylactic axillary lymph node dissections could be dramatically reduced. However, the predictive power of available standard parameters was found not high enough to limit the number of axillary lymph node dissections. Interestingly, among the factors identified to correlate with nodal involvement by Brenin et al. (36) was ER status, which we also found influential in the multivariate analysis, but not in a previous study (37). In this study, core biopsies were taken from invasive tumors of different sizes, 40% of which were classified to be T1 (Table 1). To restrict the data analyses to this group or exclude the largest T2/T3 tumors will be an interesting aspect in subsequent studies in which the number of samples will have to be increased substantially to gain reliable information for subgroups of tumors defined by size or other characteristics, or by menopause status of the patients. Taken into account that this feasibility study was primarily intended to show the potential of new assay technology in diagnostic approaches, the resulted parameter sets allow the prediction of nodal involvement with surprisingly high sensitivity and specificity, and better than estimations based on conventional presurgical diagnosis. Although the defined systems will have to be validated in an additional study and the achieved high sensitivity and specificity of 80% (large set of parameters) seem not yet sufficient for a clinical decision against axillary lymph node dissection in all cases, it can be used to recognize at least a subgroup of patients for which an N0 status is very unlikely. Furthermore, the defined core sets of few parameters will be a basis for subsequent studies with additional available assays, which will most likely lead to parameter sets with sufficiently high predicting value, meeting all requirements of clinical routine diagnosis. Consequently further methodologic refinement of multiparametric protein analyses and associated clinical studies will most likely lead to parameter sets that can be used to differentiate between different subtypes of breast cancer similar to a subclassification according to RNA expression signatures. In this pilot study, however, the numbers of the rarer histologic subtypes were too low to allow such analysis with sufficient statistical power.

**Disclosure of Potential Conflicts of Interest**

Thomas Jooss is a member of the Scientific Advisory Board of Luminex Corp. and received research support from Luminex Corp. and Merck KGaA, Darmstadt, Germany. No other potential conflicts of interest were disclosed.

**References**

Prediction of Nodal Involvement in Breast Cancer Based on Multiparametric Protein Analyses from Preoperative Core Needle Biopsies of the Primary Lesion

Georg Sauer, Nicole Schneiderhan-Marra, Cornelia Kazmaier, et al.


Updated version Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/14/11/3345

Cited articles This article cites 37 articles, 16 of which you can access for free at: http://clincancerres.aacrjournals.org/content/14/11/3345.full#ref-list-1

Citing articles This article has been cited by 3 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/14/11/3345.full#related-urls

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.