Comparison of HER2/neu Status Assessed by Quantitative Polymerase Chain Reaction and Immunohistochemistry

Frances P. O’Malley, MB,1,2 Robert Parkes, MSc,3 Eleanor Latta, MD,1,2 Suzanna Tjan, MLT,1 Tanya Zadro, BSc,4 Rosemary Mueller, PhD,4 Nona Arneson, BSc,4 Martin Blackstein, MD,5 and Irene Andrulis, PhD1,2,4,6

Key Words: Breast cancer; HER2/neu; c-erbB-2; Prognostic factor; Predictive factor; Immunohistochemistry

Abstract

We prospectively evaluated a series of 254 breast cancers by quantitative polymerase chain reaction (PCR) and immunohistochemistry using 3 antibodies: HercepTest, CB11, and TAB250. DNA was extracted from a 10-µm tumor section for PCR, and 4-µm serial sections were taken from the same block for immunohistochemistry. The immunohistochemical results were scored using a semiquantitative immunohistochemical system. A positive tumor by immunohistochemistry had a score of 5 or more. The manufacturer’s recommended scoring system was used for the HercepTest. Tumors were positive for gene amplification if the ratio of the HER2/neu gene to control gene after normalization was 2 or more. Of 254 cases, 61 showed gene amplification. For immunohistochemistry, 23% of tumors were positive with CB11, 27% with TAB250, and 37% with the HercepTest. Results for each antibody were compared with PCR results. The overall concordance for the HercepTest was 82%, which was significantly lower than that for CB11 (88%) or TAB250 (87%). The specificity for the HercepTest was 80% compared with 90% for TAB250 and 93% for CB11, while the positive predictive value for the HercepTest was 57% compared with 71% and 76% for TAB250 and CB11, respectively.

The neu gene was first identified in the DNA of chemically induced rat neuroblastomas.1 This gene has homology with the epidermal growth factor receptor and codes for a 185-kd transmembrane oncoprotein.2 The human equivalent of neu was cloned independently from a complementary DNA library and called “HER-2.”3 It also was cloned from genomic DNA and referred to as “c-erbB-2.”4 HER2/neu gene amplification, which usually results in overexpression of the encoded transmembrane protein p185, occurs in about one third of breast cancers.5-7

Slamon and colleagues5 were the first to show an association between HER2/neu gene amplification and prognosis. They demonstrated a statistically significant correlation with decreased disease-free survival and overall survival in a subset of 86 patients with lymph node–positive disease, by both univariate and multivariate analysis. This result was confirmed in a larger series of patients with lymph node–positive disease by Slamon et al6 and a number of other groups.7-15

More recently, HER2/neu gene amplification has been associated with decreased disease-free survival and overall survival in patients with lymph node–negative disease.16,17 In the study by Press et al,16 HER2/neu gene amplification was assessed by fluorescence in situ hybridization (FISH) in a series of 242 untreated patients with lymph node–negative breast cancer. In these patients, the relative risks of early recurrence, any recurrence, and disease-related death were statistically significantly associated with HER2/neu amplification by FISH. Although the results strongly suggest that HER2/neu amplification by FISH may provide independent prognostic information, this cohort may not be representative of patients with lymph node–negative disease in general. This is suggested by the fact that the traditional prognostic factors...
of tumor size and histologic grade showed no prognostic significance in this series even by univariate analysis.\textsuperscript{16}

The study by Andrulis et al\textsuperscript{17} is the largest study to date that prospectively evaluated HER2/neu amplification in tumors from patients with lymph node–negative disease. Amplification status was assessed by quantitative polymerase chain reaction (PCR) performed on DNA isolated from frozen tumors from a consecutive series of 580 lymph node–negative breast cancer cases.\textsuperscript{17} HER2/neu amplification was noted in 20\% of cases and was statistically significantly (\(P = .002\)) associated with increased risk of disease recurrence. The traditional prognostic markers also showed statistical significance in univariate analysis but not in multivariate analysis in this series of 580 patients.

Quantitative PCR is a relatively costly and labor-intensive method of assessing HER2/neu amplification. Immunohistochemistry, on the other hand, is much more readily applicable to the clinical setting. The purpose of the present study was to compare the results obtained by immunohistochemistry using 3 commonly used immunohistochemical antibodies with the results obtained by quantitative PCR.

## Materials and Methods

Consecutive cases of invasive breast cancer received for evaluation of HER2/neu accessioned at Mount Sinai Hospital, Toronto, Ontario, were included in the study. The majority of these tumors were from patients with metastatic carcinoma who were being considered for trastuzumab (Herceptin) therapy. A representative section of tumor was selected from each case and three 4-\(\mu\)m sections were taken from the block for immunohistochemical studies. A further 10-\(\mu\)m section was taken from the block for DNA extraction and quantitative PCR.

### Immunohistochemical Studies

Immunohistochemical studies were performed using the HercepTest (DAKO, Carpinteria, CA), a kit containing a prediluted polyclonal antibody A0485; TAB250 (Zymed, San Francisco, CA), a monoclonal antibody to the extracellular domain of the protein; and CB11 (Novocastra, Newcastle upon Tyne, England), a monoclonal antibody to the intracellular domain of the protein.

The 4-\(\mu\)m tissue sections were deparaffinized in two 5-minute changes of xylene and were rehydrated through alcohols to distilled water. Nonspecific reactivity was blocked with 3\% hydrogen peroxide and 5\% normal serum in phosphate-buffered saline. The TAB250 antibody was applied at a dilution of 1:150 for 1 hour at room temperature after the sections had been digested with the protease ficin (Digest-All 1, Zymed) for 10 to 15 minutes at 37\(^\circ\)C. The CB11 antibody was applied at a 1:200 dilution for 1 hour at room temperature. No heat-induced epitope retrieval method or digestion solution was applied to the slides before the application of CB11. After several washes in tris(hydroxymethyl)-aminomethane (Tris)-buffered saline at pH 7.6, sections were incubated for 30 minutes at room temperature with biotinylated goat antimouse immunoglobulin (ID Labs, London, Ontario) followed by another 30-minute incubation with streptavidin–horseradish peroxidase (ID Labs). 3,3’-Diaminobenzidine (Sigma Chemicals, St Louis, MO) was used as the chromogen, and the sections were counterstained with hematoxylin. Positive tissue controls were included in each run and consisted of sections of breast tumor known to overexpress HER2/neu protein and to show HER2/neu gene amplification. Negative controls consisted of substituting Tris-buffered saline for HER2/neu antibody.

The manufacturer’s recommended method was followed when using the HercepTest kit. The antigen retrieval method used in this kit involved immersing slides in a preheated (95\(^\circ\)C) DAKO Epitope Retrieval Solution followed by heating in a water bath at 95\(^\circ\)C for 40 minutes. This was followed by a 20-minute cool-down period at room temperature. Slides were incubated with the primary antibody (prediluted rabbit polyclonal antibody, A0485) for 30 minutes at room temperature. The slides were then incubated with the DAKO Visualization Reagent (dextran polymer conjugated with horseradish peroxidase and goat antirabbit immunoglobulins) for 30 minutes. Diaminobenzidine was used as the chromogen, and the slides were counterstained with hematoxylin.

Only membrane staining of tumor cells was considered in the analysis of the immunohistochemical results. The immunohistochemical staining was assessed using a semi-quantitative scoring system developed by Allred et al\textsuperscript{20,21} and used by O’Malley et al\textsuperscript{20} and Tuck et al.\textsuperscript{21} Briefly, this scoring system is the sum of a proportion score and an intensity score. The proportion score is an estimate of the proportion of positive cells on the entire slide and is divided into the following categories: 0, no cells stained; 1, less than 1%; 2, 1\% to 10\%; 3, 11\% to 33\%; 4, 34\% to 67\%; and 5, more than 67\%. The intensity score estimates the average staining intensity of positive tumor cells: 0, negative; 1, weak; 2, moderate; and 3, intense. The combined score ranges from 0 to 8. Before analysis of the data, a combined score of 5 or more was considered positive for HER2/neu protein overexpression. In addition, the scoring system recommended for use with the HercepTest was used when evaluating staining with this antibody: 0, no staining or less than 10\% membrane staining; 1+, partial membrane staining in any percentage; 2+, weak or moderate complete membrane staining in more than 10\% of cells; 3+, moderate or strong, complete membrane staining in more than 10\% of cells.
DNA Extraction and Quantitative PCR

Andrulis et al\textsuperscript{17} analyzed the neu DNA copy number from flash-frozen specimens. However, since the majority of cases were metastatic in the present study, frozen tissue was not available, and we used formalin-fixed paraffin-embedded specimens to determine the HER2/neu DNA copy number. We modified our procedure for quantitative PCR from archival material and used primers that produce smaller products, as described by Neubauer et al.\textsuperscript{22} Unstained 10-µm sections were obtained from formalin-fixed paraffin-embedded blocks, and an area of invasive carcinoma was marked on a corresponding H&E-stained surgical pathology slide. The unstained sections were deparaffinized by washing in xylene and consecutively increasing dilutions of ethanol. The slides were stained with hematoxylin for 30 seconds, in xylene and consecutively increasing dilutions of ethanol. Gels then were stained in 5 ng/mL of ethidium bromide in water for approximately 20 minutes and destained for 20 minutes in water. Gels then were photographed with an MP6 Landcamera (Polaroid Corp, Toronto, Ontario) onto Polaroid 665 positive-negative film according to the manufacturer’s instructions. The negatives were analyzed using a laser densitometer.

In each experiment, T47D and Skbr3 cell lines were included as controls for unamplified and amplified (approximately 3-fold) copies, respectively. The ratio of the neu gene to the control gene (n-ras or gamma-IFN) was obtained for each replicate of the samples, and at least 2 linear values were averaged. The neu: control gene ratio for each sample was normalized to the unamplified T47D value and used as the relative amplification level of the sample in question. We compared the neu DNA copy number obtained with a previous method\textsuperscript{17} and with the method described in this section for the control cell lines and for a group of 8 frozen specimens. We observed complete concordance of the values for the neu DNA copy number determined with the PCR primers and methods used previously\textsuperscript{17} on frozen specimens and the primers and methods used in the present study for paraffin specimens. A relative increase in the DNA copy number of 2-fold or more was considered amplified.

Statistical Methods

Sensitivity, specificity, concordance, negative predictive value, positive predictive value, and kappa were calculated for each antibody, using PCR results as the reference. The McNemar test was used to test for differences between HercepTest and each of the other antibodies with respect to sensitivity, specificity, and concordance. The Wilcoxon rank sum test was used to test the association between HER2/neu amplification and tumor grade, estrogen receptor (ER) and progesterone receptor (PR) status, using stratification to control for other variables where stated. Receiver operating characteristic (ROC) curves\textsuperscript{23} were constructed for each antibody by plotting their calculated true-positive rate (sensitivity) against the corresponding false-positive rate (1 minus specificity) for each possible cutoff in the ordinal scoring system used for that antibody. Optimality of the chosen cutoff was assessed by comparing these curves with reference lines representing equal trade-off between sensitivity and specificity under appropriate assumptions.

Results

Histopathologic Features of Tumors

Of 254 cases, 231 (90.9\%) were infiltrating carcinomas of no special type (ductal). There were 20 infiltrating lobular carcinomas, 1 mucinous carcinoma, 1 adenoid cystic carcinoma, and 1 metaplastic carcinoma.

By using the Nottingham modification of the Bloom-Richardson grading system,\textsuperscript{24} 25 tumors (9.8\%) were grade I/III, 92 (36.2\%) were grade II/III, and 137 (53.9\%) were grade III/III. The metaplastic carcinoma was not graded. There were 49 lymph node–negative tumors and 45 lymph
node–positive tumors, and the nodal status was unavailable in 13 cases. Distant metastases had developed in 147 patients. ER status was available for 216 cases. In 151 cases (69.9%), the tumor was positive for ER, and in 118 (55.4%) of 213 cases, the tumor was positive for PR.

**HER2/neu Status by Quantitative PCR**

All cases but one were evaluable by quantitative PCR. At least 2-fold amplification of the HER2/neu gene was shown in 61 (24.1%) of the tumors. Three cases showed borderline amplification. There was a statistically significant difference in HER2/neu amplification with respect to grade; 2 (8.0%) of 25 grade I/III tumors showed HER2/neu amplification, 15 (16.3%) of 92 grade II/III tumors were positive, and 43 (31.6%) of 136 grade III/III tumors showed HER2/neu amplification ($P = .001$).

There was also a statistically significant difference in HER2/neu amplification and hormone receptor status. Only 25 (16.6%) of 151 ER-positive tumors demonstrated HER2/neu amplification compared with 28 (43.8%) of 64 ER-negative tumors ($P < .0001$). This association seems to be independent of tumor grade. Only 18 (15.3%) of 118 PR-positive tumors demonstrated HER2/neu amplification compared with 34 (36.2%) of 94 PR-negative tumors ($P = .0005$); however, the association was not significant when ER status was controlled for.

**HER2/neu Status by Immunohistochemical Studies**

By using the HercepTest, 85 (36.6%) of the 232 tumors were positive for HER2/neu, 47 (20.3%) scored 2+, and 38 (16.4%) scored 3+. By using the semiquantitative immunohistochemical scoring system, 84 (36.2%) of 232 scored 5 or more. The rate of HER2/neu overexpression as assessed by immunohistochemistry compared with the TAB250 antibody was 16 (27.2%) of 254 and that with CB11 was 58 (22.8%) of 254. The mean immunohistochemical score for the cases positive with both TAB250 and CB11 was 6.7 (Figure 1).

The overall concordance of the HercepTest with quantitative PCR was 82%, while that for TAB250 and CB11 was 87% and 88%, respectively. While the sensitivity was high for the HercepTest (89%) compared with TAB250 and CB11 (80% and 72%, respectively), the specificity of 80% was significantly lower than that for TAB250 (90%, $P = .0004$) and CB11 (93%, $P = .00001$). Correspondingly, the positive predictive value for the HercepTest was much lower than that for TAB250 or CB11 (57% HercepTest vs 71% TAB250 and 76% CB11) (Table 1).

The sensitivities and specificities for each cutoff were determined for all 3 antibodies, and this information was used to construct ROC curves to determine the optimal cutoff for each antibody. The scores from the scoring system recommended for use with the HercepTest also were plotted.

The ROC curves for the 3 antibodies are presented in Figure 2. The 2 reference lines represent lines of equal trade-off between sensitivity and specificity when the prevalence of the HER2/neu gene is 24% (A) or 44% (B). The prevalence (based on PCR) of HER2/neu was 24% in the overall sample of 254 cases and 44% in the cases with negative ER status. Reference line B is relevant only if one has previous knowledge that the ER status is negative. For both reference lines, it was assumed that the relative “costs” of false-negative and false-positive results were equal. The ROC curves for TAB250 and CB11 are almost identical, while that for HercepTest shows greater sensitivity but lower specificity. In particular, the HercepTest ROC curve is farther from both reference lines than either TAB250 or CB11, indicating that it is not superior to the other antibodies unless the cost of a false-negative result is much greater than that of a false-positive result. For TAB250 and CB11, the optimal cutoff is 6 on the semiquantitative scale used for those antibodies, rather than a score of 5. However, the improvement is modest; 227 and 230 of 254 cases would have been classified correctly using 6 as the cutoff for TAB250 and CB11, respectively, or 5 and 7, respectively, more than obtained using 5 as the cutoff. For ER-negative cases, 5 or 6 could be used as the cutoff, since both values are almost the same distance from B.

**Discordant Cases**

There were 8 cases in which all the antibodies were negative, but quantitative PCR showed gene amplification. The level of gene amplification was low, however, with all cases showing less than 4-fold amplification.

In 11 cases, all antibodies showed HER2/neu overexpression, but the quantitative PCR was negative for gene amplification. Of the cases stained with TAB250 and CB11, the mean score was 5.9, which was near the cutoff of 5. Of the cases stained with the HercepTest, 5 scored 2+ and 6 scored 3+.

**Discussion**

While there is a consensus with respect to the prognostic significance of HER2/neu overexpression in lymph node–positive disease,5–15 the prognostic significance of HER2/neu alterations in lymph node–negative disease is more controversial, particularly HER2/neu overexpression, as assessed by immunohistochemistry.25–38 There are possible technical explanations for these conflicting results, including use of different antibodies, different fixatives, and variation in interpretation of staining.

A comparison of different antibodies was performed several years ago by Press et al.39 They analyzed immunohistochemical staining with 28 antibodies in a series of breast
tumors that had HER2/neu amplification assessed by Southern blot and RNA and protein levels assessed by Northern and Western blots, respectively. Each of the antibodies used detected some of the breast tumors with known HER2/neu overexpression, but none detected all the samples known to overexpress HER2/neu. Only immunostaining of the cell membranes of breast cancer cells correlated with HER2/neu overexpression. Several of the other antibodies showed nonspecific cytoplasmic staining. This cytoplasmic staining may have been scored as positive in some studies,
possibly contributing to some of the discrepant results with respect to the prognostic significance of HER2/neu reported in the literature.39 The study by Press et al,39 however, is not directly applicable to the immunohistochemical methods that are used today, as antigen retrieval methods were used only with one of the antibodies studied, whereas antigen retrieval techniques are recommended with a number of the commercially available antibodies currently in use. Also, the peroxidase-antiperoxidase detection method was used in the study by Press et al,39 yet this has been largely substituted in current practice by the more sensitive avidin-biotin complex systems. Nevertheless, there were 8 cases in our study in which there was no staining with any of the 3 antibodies, although quantitative PCR showed amplification, albeit at a low level. This suggests that even the most sensitive antibodies and detection systems may fail to detect protein overexpression when associated with low levels of amplification.

The type of tissue fixation used also influences immunohistochemical staining. Four fixatives were compared (Bouin fluid, buffered formalin, alcoholic formalin, and methacarn),40 and the most consistent results were obtained with the antibody CB11 on frozen tissues and with the TAB250 clone on paraffin-embedded sections, especially if buffered formalin had been used.

Different scoring systems and cutoffs have been used in assessment of HER2/neu immunohistochemical studies, and this probably also has contributed to some of the discrepancies between results. In many studies, the immunohistochemical results typically are scored as positive or negative. The cutoff for a positive result can vary widely, however. In the study by Paik et al,41 a tumor was scored positive if any tumor cell showed definite membrane staining. In the Cancer and Leukemia Group B protocol 8541,42 the percentage of HER-2/neu staining was recorded and divided into low HER2/neu expression (< 50% membrane positivity) and high HER-2/neu expression (50% or greater membrane positivity). The scoring system developed by Allred et al18 records the percentage positivity and the intensity of staining (see “Immunohistochemical Studies” in the “Materials and Methods” section). This scoring system was used in a subset of the patients in the study by Thor et al42 and showed significant correlation with the scoring system used by Thor et al, as well as a significant correlation with response to high-dose therapy.

Patients’ eligibility for inclusion in a multicenter, randomized, controlled trial assessing chemotherapy alone or in combination with trastuzumab, was based on HER2/neu immunohistochemical staining of their tumor specimens.43 Tumor tissue was tested with an immunohistochemical assay referred to as the Clinical Trial Assay. Tumors were scored as 0, 1+, 2+, or 3+, and only patients with 2+ or 3+ staining were considered eligible for the study (approximately 33% of patients). Results indicated that the beneficial effects of treatment were confined to the group of patients whose

![Table 1](image)

<table>
<thead>
<tr>
<th></th>
<th>HercepTest</th>
<th>TAB250</th>
<th>CB11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>80</td>
<td>90†</td>
<td>93†</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89</td>
<td>80</td>
<td>72†</td>
</tr>
<tr>
<td>Concordance</td>
<td>82</td>
<td>87†</td>
<td>88†</td>
</tr>
<tr>
<td>Negative PV</td>
<td>96</td>
<td>94</td>
<td>91</td>
</tr>
<tr>
<td>Positive PV</td>
<td>57</td>
<td>71</td>
<td>76</td>
</tr>
<tr>
<td>kappa</td>
<td>57</td>
<td>67</td>
<td>66</td>
</tr>
</tbody>
</table>

PV, predictive value.
† Values are given as a percent.
‡ Significantly different from HercepTest (P < .05).

![Figure 2](image)

Receiver operating characteristic curves for HercepTest (DAKO, Carpinteria, CA), TAB250 (Zymed, San Francisco, CA), and CB11 (Novocastra, Newcastle upon Tyne, England) based on 254 cases of invasive breast cancer with quantitative polymerase chain reaction results used as reference. True-positive and false-positive rates are plotted together using each point as indicated on the semiquantitative scale as the cut point. TAB250 and CB11 results were almost identical, while the HercepTest demonstrated higher sensitivity and lower specificity relative to the other 2 antibodies. Reference lines are used to demonstrate points of equal trade-off between sensitivity and specificity based on the assumption of HER2/neu prevalence of 24% (A) or 44% (B). For neither reference line was the HercepTest superior to either CB11 or TAB250. For reference A, corresponding to the usual case in which the estrogen receptor (ER) status is unknown, it seems that 6 may be a somewhat better cutoff than 5 for both TAB250 and CB11. Reference B corresponds to the case in which ER status is known to be negative.
tumors demonstrated the strongest (3+) Her2/neu immuno-
staining.43,44

The US Food and Drug Administration and the Health Protection Branch of Canada approved the use of the DAKO HercepTest for the immunohistochemical detection of HER2/neu overexpression in tumors of patients who are being considered for trastuzumab therapy. This test has not been studied directly in a clinical trial; however, it has been compared with the Clinical Trial Assay in more than 500 breast tumor samples obtained from the National Cancer Institute Breast Cancer Tissue Bank.44 Results indicated a 79% concordance rate between the 2 tests; however, 42% of the cases showing 2+ staining with the HercepTest were negative with the Clinical Trial Assay.44 A further study reported a specificity of only 42% with the HercepTest when evaluated in a group of 48 cases that previously had been found negative when assessed by immunohistochemistry using the same antibody as used in the HercepTest (A0485) but a different detection kit, as well as by FISH.45 In the present study, the HercepTest was associated with the lowest specificity (Table 1). Also, of 11 cases that were positive by all 3 antibodies but did not show amplification, 6 showed 3+ staining with the HercepTest, while the mean score for the other 2 antibodies was near the cutoff (5.9). Although the antibodies TAB250 and CB11 showed lower sensitivities than the HercepTest, these antibodies exhibited higher specificities and positive predictive values and, therefore, may be more clinically useful as prognostic and predictive indicators.

Given the widespread use of immunohistochemical tests in the clinical setting, there is a need to address the issues of limited standardization of the technical procedures and scoring methods used for immunohistochemical tests for HER2/neu.

From the 1Department of Pathology and Laboratory Medicine, Mount Sinai Hospital; 2Department of Laboratory Medicine and Pathobiology, University of Toronto; 3Division of Epidemiology and Biostatistics and 4Centre for Cancer Genetics, Samuel Lunenfeld Research Institute; 5Department of Medicine, Mount Sinai Hospital and University of Toronto; and 6Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario.

This study was supported by a grant from Hoffmann-La Roche, Mississauga, Ontario.

Address reprint requests to Dr O’Malley: Mount Sinai Hospital, Dept of Pathology and Laboratory Medicine, 600 University Ave, Toronto, ON, M5G1X5 Canada.

References


