Breast Cancer Research and Treatment (2005) 93: 3–11
DOI 10.1007/s10549-004-6275-8

Report

Determination of HER2 gene amplification by fluorescence *in situ* hybridization and concordance with the clinical trials immunohistochemical assay in women with metastatic breast cancer evaluated for treatment with trastuzumab

Noël Dybdal[1], Grazyna Leiberman[1], Steven Anderson[2], Bryan McCune[2], Alex Bajamonde[1], Robert L. Cohen[1], Robert D. Mass[1], Corsee Sanders[1], and Michael F. Press[3]
[1]*Genentech Inc., South San Francisco, CA;* [2]*Laboratory Corporation of America, Research Triangle Park, NC;* [3]*Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA*

Key words: fluorescence *in situ* hybridization (FISH), HER2, immunohistochemistry (IHC), metastatic breast cancer (MBC), trastuzumab

Summary

Purpose. To evaluate the concordance between HER2 gene amplification, determined by fluorescence *in situ* hybridization (FISH), and HER2 protein overexpression assessed by an immunohistochemical (IHC) assay. The IHC protocol used was a research assay, known as the Clinical Trial Assay (CTA), developed to select women with metastatic breast cancer (MBC) for three pivotal clinical trials of trastuzumab therapy.

Methods. A direct-labeled, dual-probe FISH assay was used to determine HER2 amplification in 623 fixed breast cancer tissue specimens. These specimens had been stored as paraffin-embedded sections for 2–5 years. All specimens had been analyzed for HER2 protein expression by the CTA. To assess the reproducibility of FISH results in archived material, we evaluated a separate group of 617 breast cancer tissue specimens at two different laboratories.

Results. Informative FISH results were available for 529 (85%) of the 623 specimens. Overall concordance between FISH and IHC results was 82% (95% CI; 78–85%). Assay agreement between FISH results and specimens with immunostaining scores of 0, 1+, and 3+ were 97, 93 and 89%, respectively. However, only 24% of specimens with 2+ immunostaining scores had HER2 amplification by FISH; there was assay disagreement in 76% of specimens in this IHC subgroup. Interlaboratory FISH concordance was 92% (95% CI; 89–94%), indicating very good assay reproducibility in these archived specimens.

Conclusion. HER2 status determined by CTA-IHC and FISH are significantly correlated; however, differences between these two assays can affect patient selection for trastuzumab therapy.

Introduction

The human epidermal growth factor receptor type 2 (HER2) proto-oncogene, located on chromosome 17q21, encodes a 185-kDa transmembrane protein that is a member of the epidermal growth factor receptor family of tyrosine kinases [1–3]. The HER2 gene is amplified in 20–30% of human breast cancers, and amplification is directly linked to overexpression of the HER2 protein product [4–7]. In breast cancer, HER2 gene amplification and protein overexpression are associated with a poor clinical prognosis in terms of shorter disease-free and overall survival for women with node-positive or node-negative disease [4, 6, 8–11].

Trastuzumab (Herceptin, Genentech Inc., South San Francisco, CA) is a humanized monoclonal antibody directed against an epitope on the extracellular domain of the HER2 protein. In preclinical studies, trastuzumab was most effective against tumor cells with HER2 amplification and protein overexpression [12, 13]. The eligibility criteria for the three initial pivotal clinical trials of trastuzumab in women with metastatic breast cancer (MBC) required documentation of HER2 protein overexpression [14–16]. A research immunohistochemical (IHC) assay (the Clinical Trial Assay [CTA]) was developed to help select patients for entry to these pivotal clinical trials. The CTA involved two murine monoclonal antibodies, 4D5 (the murine precursor of trastuzumab) and CB11, each directed against different epitopes on the HER2 protein. The assay was subjectively evaluated for HER2 protein expression using a scoring scale of 0, 1+, 2+, or 3+ to reflect increasing degrees of protein immunostaining. Only patients with tumors having HER2 immunostaining scores of 2+ or 3+ were eligible for the clinical trials.

While its commercial development was considered impractical, the CTA served as a prototype assay for two other IHC assays for HER2 that have Food and Drug Administration (FDA) approval: the DAKO HercepTest (DAKO Corporation, Carpinteria, CA) and the Ventana Pathway (Ventana Medical Systems, Tucson, AZ) assay. The DAKO HercepTest was compared

directly with the CTA to demonstrate concordance of greater than 75% and obtained FDA approval. The Ventana Pathway assay uses one of the antibodies, CB11, specified in the CTA, and was compared with the FDA-approved DAKO HercepTest to show concordance. Clinicians now use these IHC assays routinely to select patients for trastuzumab therapy. However, there is still much controversy over the reliability of IHC assays to accurately assess HER2 protein expression in routine clinical specimens [11, 17–34]. It is well known that a number of factors can affect IHC assay results, including tissue fixation and processing, reagent variability, interpretation, and scoring. For example, IHC assays are much less able to accurately determine HER2 protein overexpression in paraffin-embedded tissue specimens than in fresh frozen tissues specimens [6].

The FDA has also approved assays using fluorescence *in situ* hybridization (FISH) for HER2 assessment in clinical samples. These assays directly assess HER2 gene amplification. In a study comparing FISH and IHC assays for the detection of HER2 in a group of molecularly characterized breast cancer tissue specimens, FISH was the most sensitive and accurate assay method currently approved by the FDA for HER2 evaluation in clinical specimens [35]. While there are a number of published papers [23, 35, 36] comparing FDA-approved IHC assays (DAKO HercepTest and Ventana Pathway assay) with FISH assays, there are no published comparisons of the CTA (used to screen women for entry to the trastuzumab pivotal trials) and FISH.

In this study, we evaluated the concordance between HER2 gene amplification determined by FISH and HER2 protein overexpression previously determined by IHC in breast cancer tissue specimens from women screened for three pivotal clinical trials of trastuzumab. Because the samples used for the FISH analysis had been archived for several years, we also assessed the reproducibility of FISH results by evaluating a series of specimens at two different laboratories.

Methods

Tissue specimens

As part of the eligibility screening for the three initial pivotal clinical trials of trastuzumab in MBC, 5998 tumor tissue specimens were submitted to a central reference laboratory (Laboratory Corporation of America, Research Triangle Park, NC) (Laboratory 1). Ten to 12 sections (4 to 6 μM) were cut from each tissue specimen and mounted on positively charged slides. One section was stained with hematoxylin and eosin to confirm the presence of a tumor. Remaining sections were either used for IHC assays to determine HER2 protein expression using the CTA (methods below) and estrogen- and progesterone-receptor status, or archived. Assay results for each tissue specimen were recorded, and all processed and unused sections were archived at room temperature.

Specimen selection and statistical analysis

Specimens were first reviewed to identify those with at least two remaining unstained tissue sections. Based on previous immunostaining scores, specimens were divided into two groups: those with immunostaining scores of 0 or 1+ (considered negative for HER2 protein overexpression) and those with immunostaining scores of 2+ or 3+ (considered positive for HER2 protein overexpression). Approximately 300 specimens from each group were randomly selected to determine HER2 amplification using FISH (methods below). These FISH analyses were performed at Laboratory 1, with investigators blinded to the previous CTA scores.

Using the premise that assay concordance of less than or equal to 75% was unacceptable, it was estimated that a sample size of 600 specimens would provide 90% power to detect a level of concordance 5% higher than the unacceptable level at a 5% significance level using a one-sided test of proportion. Concordance was defined as the proportion of specimens with CTA immunostaining scores of 0 or 1+ and negative for HER2 gene amplification by FISH, plus the proportion of specimens with CTA immunostaining scores of 2+ or 3+ and positive for HER2 gene amplification by FISH. Overall assay concordance was calculated and the value of the κ statistic estimated [37].

Fluorescence in situ hybridization (FISH) concordance between laboratories

To assess possible variability in FISH results between two laboratories, a separate group of 617 archived breast cancer tissue specimens were analyzed both at Laboratory 1 and at the University of Southern California (USC, Los Angeles, CA) (Laboratory 2). The original paraffin blocks from which the specimens were cut were no longer available. Laboratory 1 performed FISH analysis using excess 4 to 6 μM sections collected from the blocks at the time of the original screening. These sections were stored at ambient temperature on glass slides in their paraffinized state. Laboratory 2 performed FISH analysis using previously immunostained tissue sections (methods below). These sections had been used as negative controls for the CTA or estrogen and progesterone receptor IHC assays and had been stored on coverslipped slides at ambient temperature. Concordance between FISH results from the two laboratories was calculated and the κ statistic was evaluated. Only samples that yielded informative results at both laboratories were included in the concordance analysis.

Clinical Trials Assay (CTA)

The CTA was first developed as a single IHC analysis using the monoclonal antibody 4D5 (Genentech, Inc., South San Francisco, CA). This antibody is the murine precursor of trastuzumab and binds to an extracellular epitope of the HER2 protein near the surface of the cell

membrane. In response to data suggesting that 4D5 is a potentially sub-optimal IHC reagent when used in protocols without antigen retrieval [19], a second murine monoclonal antibody, CB11 (Novocastra Laboratories, Ltd., Newcastle upon Tyne, UK), was added to the CTA. The CB11 antibody binds to an intracellular epitope near the C-terminus of the HER2 protein. The 4D5 and CB11 IHC assays differ in antigen retrieval methods, but both use a similar visualization method involving a biotinylated horse anti-mouse antibody and a standard avidin-biotin horseradish peroxidase complex.

All analyses for the CTA were performed at Laboratory 1; staining procedures were done with a TechMate 1000 autostainer (Biotek Solutions, Inc., Santa Barbara, CA). In brief, tissue sections were deparaffinized through xylene-graded alcohols and rinsed. For the 4D5 assay, a hydrogen peroxide blocking step was followed by antigen retrieval involving limited protease digestion for 10 min at 37 °C. After rinsing, a routine serum block was performed and the sections incubated overnight with the 4D5 antibody. For visualization, sections were incubated first with a biotinylated horse anti-mouse antibody (PK-6100, Vector Laboratories, Burlingame, CA), followed by incubation with a standard avidin-biotin horseradish peroxidase enzyme complex and the chromogen 3,3′-diaminobenzidine tetrachloride (DAB). Slides were then counterstained with Harris hematoxylin.

For the CB11 assay, antigen retrieval involved boiling deparaffinized tissue sections in pH6 citrate buffer in a microwave oven twice for 5 min each. Following a normal serum block, sections were incubated with the CB11 antibody for 25 min at room temperature and processed using the same visualization procedures as above.

A pathologist evaluated all specimens and used the same staining interpretation criteria for both 4D5 and CB11 assays. A score of 0 was assigned to specimens with no membrane staining or staining in less than 10% of tumor cells. Specimens with staining in more than 10% of tumor cells were scored as 1+ for partial membrane staining, 2+ for weak to moderate complete membrane staining, and 3+ for moderate to strong complete membrane staining. If either antibody showed positive immunostaining the case was scored according to the stronger staining reaction. The human breast cancer cell lines MDA-MB-231, MDA-MB-175, and SK-BR-3 were used as performance control standards in every assay (both 4D5 and CB11), with corresponding IHC staining scores of 0, 1+, and 3+, respectively. Patients were deemed eligible for the trial if their biopsy specimen scored 2+ or 3+ by either assay.

Fluorescence in situ hybridization (FISH)

FISH was performed at Laboratory 1 using the PathVysion assay kit (Vysis, Inc., Downers Grove, IL), which includes two directly labeled DNA probes; a locus-specific probe for the HER2 gene labeled with SpectrumOrange and an alpha satellite probe targeting the centromere region of chromosome 17 (CEP17) labeled with SpectrumGreen. The assay was performed according to the manufacturer's instructions, with minor modifications in the pretreatment and protease digestion incubation times to optimize conditions for archived tissue specimens. Assays were performed either manually or using the VP2000 automated platform (Vysis, Inc., Downers Grove, IL).

In brief, archived hematoxylin and eosin stained tissue sections were reviewed for presence and location of tumor. The tumor location was marked and a corresponding location etched on a FISH slide for post-hybridization scoring. Following deparaffinization, specimens were pretreated by incubation in 0.2N HCl for 20 min, rinsed, incubated in pretreatment solution (80 °C, 60 min), rinsed again, and then digested with protease (37 °C, 30 min). Specimens were rinsed, dried, and fixed in 10% neutral buffered formalin, followed by incubation in 70% formamide/2x standard sodium citrate (70 °C, 5 min), then dehydrated with graded alcohols and air-dried. The HER2:CEP17 probe cocktail (10–20 μl) was applied, a coverslip sealed to the slide, and the specimens hybridized overnight at 37 °C. Slides were then washed and counterstained with 4′,6-diamidino-2-phenylindole (DAPI) fluorochrome stain. Sections not evaluated immediately were stored at −20 °C.

Presence of the two probes was evaluated using appropriate band-pass filters on an epifluorescence microscope. Average copy numbers of HER2 and CEP17 from 40 tumor cell nuclei were counted and the signal ratio calculated. Specimens with a HER2:CEP17 ratio of greater than or equal to 2 were considered amplified (FISH-positive), while those with a HER2:CEP17 ratio of less than 2 were considered non-amplified (FISH-negative). For specimens that were clearly non-amplified (signal ratio less than 1.8) or clearly amplified (signal ratio of greater than 2.2) after assessment of 40 cells, there was no further counting. For specimens with borderline signal ratios of 1.8 to 2.2, 20 more cells were counted and the ratio recalculated. All results were verified by independent slide review.

Laboratory 2 also used the PathVysion assay for FISH analyses. The assay was performed on previously immunostained tissue sections using the procedures described above with the following modifications. Slide coverslips were first removed by immersion in xylene for 72–96 h, followed by an ethanol rinse. Specimens were incubated in pretreatment solution for 30 min at 80 °C and digested with protease for 10–60 min at 37 °C. The volume of the HER2:CEP17 probe cocktail (10–35 μl) was adjusted according to the coverslip size required by the specimen. Average copy numbers of HER2 and CEP17 were found by counting 60 tumor cell nuclei in all cases, and specimens with a HER2:CEP17 ratio of greater than or equal to 2 were considered positive for HER2 gene amplification.

Validation experiments (data not shown) performed by both Laboratory 1 and Laboratory 2 confirmed coverslip size and probe volume across the ranges being used should not influence FISH scores. In specimens with uninformative FISH results, efforts were made at Laboratory 2 to optimize conditions and the assay was repeated, provided tissue sections remained intact or more sections were available.

Results

Of the 5998 breast cancer tissue specimens submitted to Laboratory 1 for HER2 screening between 1994 and 1997, 5251 (87%) had at least two unstained tissue sections. The distribution of immunostaining score frequencies determined by the CTA for the entire 5998 specimens were: 0, 58%; 1+, 9%; 2+, 10%; and 3+, 23%. The 5251 tissue specimens with unstained tissue sections had a similar IHC score frequency distribution to the whole group.

A total of 623 specimens were selected for FISH analysis, with 306 (49%) specimens negative for HER2 protein overexpression (immunostaining scores of 0 or 1+) and 317 (51%) specimens positive for HER2 protein overexpression (immunostaining scores of 2+ or 3+) by prior CTA analysis. An informative FISH result was obtained in 529 (85%) of these specimens. The most common reason for non-informative results was an inability to achieve hybridization of one or both probes with consequent lack of fluorescence signals.

A total of 11 of the 529 informative cases (2%) representing the screened population had FISH ratios between 1.8 and 2.2. When 20 additional cells were counted in these 11 cases at Laboratory 1 to give a total of 60 cells evaluated, none of the cases showed a change in the amplification status. Those cases with a FISH ratio less than 2.0 after scoring 40 cells continued to have a ratio less than 2.0 after 60 cells were scored. Similarly, those cases with a FISH ratio greater than or equal to 2.0 after scoring 40 cells continued to have a ratio greater than 2.0 after scoring 60 cells.

There was a high level of concordance between HER2 gene amplification determined by FISH and HER2 protein overexpression determined by the CTA. In total, 235 FISH-negative and 9 FISH-positive specimens were among those with immunostaining scores of 0 or 1+ determined by the CTA, and 88 FISH-negative and 197 FISH-positive specimens were among those with CTA scores of 2+ or 3+. The overall concordance between FISH-negative specimens with CTA scores of 0 or 1+, combined with FISH-positive specimens with CTA scores of 2+ or 3+, was 82% (95% CI; 78–85%), significantly higher than the pre-specified unacceptable level ($p < 0.0004$). The κ statistic was 0.63 (95% CI; 0.57–0.69), indicating fair to good assay agreement.

The rate of FISH-positive specimens within each CTA score subgroup varied from 3 to 89% (Table 1). There was good assay agreement in the 3+ immunostaining subgroup; 89% of specimens were positive for HER2 amplification. Similarly, 235 (96%) of the 244 specimens negative for HER2 protein overexpression (CTA score 0 or 1+) were also FISH-negative. The greatest discrepancy was in specimens with 2+ immunostaining by the CTA; only 24% of these specimens were positive for HER2 amplification by FISH. There was assay disagreement between FISH and immunostaining in 67 (76%) of the 88 specimens with CTA scores of 2+, suggesting false positive results by the CTA.

Table 1. Rate of HER2 amplification determined by fluorescence *in situ* hybridization according to immunohistochemical score

FISH amplification	IHC score				
	0	1+	2+	3+	Total
Negative	207	28	67	21	323
Positive	7	2	21	176	206
Total	214	30	88	197	529
Rate of positive FISH amplification	3%	7%	24%	89%	39%

FISH, fluorescence *in situ* hybridization; IHC, immunohistochemistry.

Table 2. Predicted incidence of HER2 amplification in 5998 patients with metastatic breast cancer previously screened by the Clinical Trial Assay for entry to Genentech pivotal clinical trials

CTA IHC score	Observed IHC score		Observed HER2 amplification rate (By CTA IHC score from Table 1)	Predicted HER2 amplification per IHC category	
	No. of patients	Percentage of total group (%)		No. of cases	Percentage of total group showing gene amplification (%)*
0	3479	58	3%	104	1.73
1+	540	9	7%	38	0.63
2+	600	10	24%	144	2.40
3+	1379	23	89%	1227	20.46
Total	5998	100		1513	25.22

*The rate is the number of predicted cases divided by the total number of patients (5998) and expressed as a percentage.
CTA, Clinical Trials Assay; IHC, immunohistochemistry.

The study was designed to give an approximately equal distribution of specimens with positive and negative immunostaining scores for HER2 protein overexpression. To predict the incidence of HER2 gene amplification in the women with MBC screened by the CTA, we extrapolated the observed rates of HER2 amplification determined by FISH to the underlying population of 5998 women originally screened by the CTA for entry to the pivotal clinical trials (Table 2). Using the observed HER2 amplification rates, the incidence of positive HER2 amplification in this population is predicted to be 25%. Of the predicted 1513 FISH-positive patients, 1227 (81%) would be expected to have 3+ immunostaining, 144 (10%) to have 2+ immunostaining, 38 (2%) to have 1+ immunostaining, and 104 (7%) to have an immunostaining score of 0. Figure 1 shows the predicted relative rates of FISH-positive and FISH-negative scores within each immunostaining subgroup in patients screened by the CTA. The overall concordance rate between FISH and CTA results using these extrapolated data was 88% (95% CI; 85 to 91%).

To establish the level of concordance among FISH assays, a second cohort of 617 patients entered in Genentech pivotal clinical trials [38] was used and independently analyzed in both laboratories. These FISH comparison cases are not the same cases randomly selected from the 5998 women screened for entry to clinical trials to determine the CTA-FISH concordance rate. Sufficient tissue was not available from all of these for a second series of FISH assays. Informative FISH results were obtained in both laboratories for 488 of 617 cases. Seventy-eight (13%) cases were non-informative at Laboratory 1 and 51 (8%) cases were non-informative at Laboratory 2.

Laboratory 1 performed a single FISH assay attempt for each case without repeating the assay in cases that were initially unsuccessful. Laboratory 2 repeated the FISH assay for any case that initially gave a non-informative result, with trouble-shooting and individual optimization of FISH conditions for that particular case. As expected, the FISH success rate for Laboratory 2 (92%) was higher than for Laboratory 1 (87%). Among the 488 informative cases used to show concordance between FISH at the two laboratories, 7 (<2%) had a FISH ratio between 1.8 and 2.2 and the amplification status was not altered when additional tumor cells were assessed.

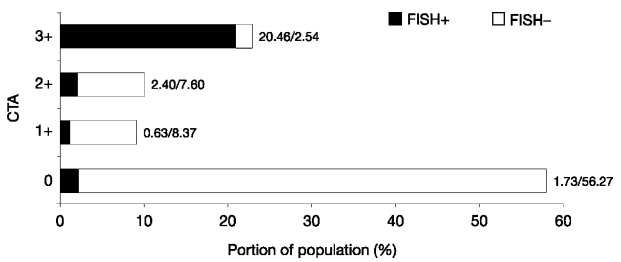

Figure 1. Extrapolated relative frequency rates of HER2 gene amplification in patients with metastatic breast cancer screened by the Clinical Trial Assay.

Table 3. Comparison of fluorescence *in situ* hybridization assay results

Laboratory 1 FISH results	Laboratory 2 FISH results		
	Negative	Positive	Total
Negative	106	37	143
Positive	3	342	345
Total	109	379	488

Overall concordance = 92% (95% CI; 89–94%); κ statistic = 0.79 (95% CI; 0.72–0.85).
FISH, fluorescence *in situ* hybridization.

The overall level of agreement between FISH results was 92% (95% CI; 89–94%), and the κ statistic was calculated to be 0.79, indicating good to excellent agreement (Table 3). There were 37 cases determined to be FISH-negative by Laboratory 1 and FISH-positive at Laboratory 2. Review of the CTA scores of these cases showed 3+ immunostaining in 35 of 37 specimens.

Discussion

Accurately determining HER2 status in clinical specimens is essential to select appropriate patients for therapy with trastuzumab. Currently, IHC and FISH are the most widely used methods for assessing HER2 status in clinical specimens [39]. The FDA has approved assays using each of these methods to help select patients for trastuzumab therapy. We found an overall level of concordance of 82% between HER2 amplification determined by FISH and IHC scores from the CTA; however, the discordance between FISH results and IHC scores has clinically important implications for trial design and selection of patients for trastuzumab therapy.

It is now appreciated that IHC assays have significant shortcomings when used to subjectively evaluate HER2 protein expression in clinical specimens. False negative or false positive results with IHC techniques can be caused by variations in specimen processing, reagents, and staining interpretation [6, 18, 19, 39–42]. In contrast, direct evaluation of HER2 gene amplification using FISH is an accurate and reliable method of evaluating HER2 gene amplification status in clinical specimens [35]. HER2 gene amplification also correlates with response to trastuzumab therapy; patients with FISH-positive MBC have higher response rates and survive longer than patients with FISH-negative MBC [38].

HER2 overexpression is closely correlated with HER2 gene amplification in frozen tissue specimens; about 98% of specimens show either non-amplified, low expression breast cancers or amplified, overexpression breast cancers [17, 23]. Therefore, assays of HER2 overexpression would be expected to identify gene amplification almost exclusively. However, we did not see this level of concordance in our study. Although there was significant agreement between FISH results

and IHC by CTA ($p < 0.0004$), the agreement rate between CTA and FISH varied by scoring category. Overall, 96% of the specimens with immunostaining scores of 0 or 1+ were negative for gene amplification. In specimens with 3+ immunostaining, 89% were positive for HER2 amplification. However, there was substantial assay disagreement among specimens with 2+ immunostaining scores; only 24% of these specimens were positive for gene amplification.

These data suggest that 76% of the specimens with 2+ immunostaining by the CTA were potentially false positive IHC results. These observations agree with a number of other reports of a high rate of FISH-negative results in samples with 2+ immunostaining using a number of IHC assays [18, 20–33]. It was speculated that a subgroup of potentially false positive IHC 2+ tumors may have true HER2 overexpression in the absence of gene amplification because of aberrant translation or transcription [6]. However, further analysis of these cases showed nearly complete agreement between the HER2 gene amplification status and expression status [17, 23].

Our study results are similar to those recently reported by Owens and colleagues who compared FISH assay results with IHC performed using the HercepTest in 6556 breast cancer specimens [43]. The rate of FISH-positive results was 4% for specimens with IHC scores of 0, 7% for IHC 1+, 23% for IHC 2+, and 92% for IHC 3+. Our results were also similar to recent data from the Breast Cancer International Research Group (BCIRG), which compared FISH results from two central reference laboratories with HercepTest or Pathway IHC assay results from outside referral laboratories [44]. The agreement rate for IHC 3+ breast cancer specimens was 79% between FISH and the HercepTest and 78% between FISH and the Pathway assay.

Because our specimens were potentially suboptimal, (having been archived for 2–5 years) the reliability of FISH results was evaluated by analyzing 617 samples by FISH at two laboratories; 488 gave informative results at both laboratories. The concordance rate between laboratories was 92%, indicating that FISH results were reproducible in these archived specimens, even when they had been previously processed using IHC staining. Data from the BCIRG also indicate that FISH results are highly reproducible, as the group observed an assay agreement rate of 92% between FISH results from its central laboratory and outside laboratories [44]. Primary tumor tissue was used to analyze HER2 status in the Genentech pivotal clinical trials, but metastatic disease was the therapeutic target for treatment. This was considered the correct strategy because of tissue availability of metastatic tumors and the highly stable nature of HER2 status throughout the progression of invasive breast cancer. Several investigators have shown that the overexpression of HER2 is unchanged in most tumors when comparing primary and metastatic sites in the same patient [45–48].

The data from this study were extrapolated to predict the expected incidence of HER2 amplification among women with MBC screened by the CTA for the Genentech pivotal clinical trials of trastuzumab. The incidence of HER2 gene amplification in these patients was estimated to be 25%, consistent with previous reports of HER2 amplification rates in breast cancer tissue specimens [4, 6, 7]. The 25% expected incidence rate of HER2 amplified specimens is also in agreement with recent reports of HER2 amplification rates by FISH in large cohorts of breast cancer specimens. The incidence of HER2 positive specimens was 23% in over 6500 cases evaluated by Owens and colleagues, and 26% in over 2500 breast cancer specimens screened by the BCIRG central laboratory for clinical trials of adjuvant trastuzumab [43, 44].

Although the concordance data from this study indicated that most of the patients with 0, 1+ and 3+ immunostaining by the CTA would also be similarly classified as 'negative' or 'positive' by FISH (Table 2, Figure 1), the disagreements in these immunostaining categories is important to note. Based on the predicted HER2 gene amplification rate, 2.36% of all breast cancers screened by the CTA would be expected to have immunostaining scores of 0 or 1+, but, nevertheless, have HER2 gene amplification by FISH. These patients would account for about 9.44% of all women with HER2-amplified breast cancers. This is an important group of women that would be denied treatment with trastuzumab, either in clinical trials or in clinical practice. The disagreements between CTA and FISH in the 3+ category are also interesting. If 3+ immunostaining were used as a criterion for treatment with trastuzumab, then 12% of women treated would be FISH-negative and unlikely to respond to therapy [16, 38]. The CTA was performed and interpreted at a single centralized laboratory. If in other clinical settings these IHC assays were performed in multiple laboratories (as is common) the proportion of women with IHC 3+, FISH-negative biopsies might be expected to be higher due to inter-laboratory variability. More disagreement between an IHC score of 3+ relative to FISH status, when the results are obtained from multiple local laboratories, has been reported in the B-31 trial of the National Surgical Adjuvant Breast and Bowel Project. In the B-31 trial the rate of HER2 false positives decreased from 21% at local laboratories to 2% when the analyzes were performed at a central laboratory [49].

Clinical diagnostic assays should satisfy a number of criteria, particularly test accuracy and test reproducibility. Concordance between assay methods does not necessarily ensure their accuracy [50]. Accuracy is the ability of a clinical test to provide results giving the 'true' status of the sample. Detection of HER2 overexpression/amplification in clinical specimens has already been shown to be less accurate for FDA-approved IHC assay methods than for FDA-approved FISH assays [23,35]. Hence, FISH is expected to provide results on HER2 status that are more similar to the actual HER2 status observed in frozen tissue samples using other methods than would be expected from IHC [23, 35]. If this is correct, the results observed here indicate that a

substantial proportion of the women whose breast cancers were screened as HER2-positive by the CTA and were, therefore, candidates for entry to trastuzumab pivotal clinical trials did not, in fact, have HER2 gene amplification. Because women with false-positive HER-2 status are not expected to respond to trastuzumab, their inclusion in these clinical trials would be expected to result in an underestimation of the response rate and, therefore, efficacy of trastuzumab.

Reproducibility is the ability of a test result to be reproduced either in the hands of other investigators or in other laboratories. Because of variation in tissue processing methods, IHC staining methods, and subjective interpretation of test results, IHC is recognized as difficult to reproduce when quantitative or semi-quantitative results are required. Detection of HER2 overexpression by IHC assays has shown a relatively low level of reproducibility between clinical laboratories [51–53]. The College of American Pathologists (CAP), through the CAP survey of clinical testing methods, found considerable variability in IHC assay results for HER2 status and a remarkably high level of reproducibility for FISH assay results [51]. This higher level of reproducibility for FISH is consistent when we compare the FISH concordance results reported here for Laboratory 1 to those from Laboratory 2.

Currently, there are several large clinical trials investigating trastuzumab for breast cancer in both adjuvant and metastatic settings. Our data suggest that FISH may identify a significant number of patients who could be eligible for trastuzumab who would not otherwise be selected due to negative IHC assay results. Our results also suggest that FISH would exclude a significant number of women from clinical trials in both the IHC 2+ and IHC 3+ categories whose breast cancers do not have HER2 gene amplification. Including these patients could lead to an underestimation of the efficacy of trastuzumab. Overall, while we found a statistically significant level of concordance between HER2 amplification determined by FISH and the HER2 protein overexpression determined by the CTA, the disagreements between these assay methods are worth considering when designing a clinical trial, to minimize both the false-negative exclusions and false-positive inclusions in a study population. These results suggest that although IHC is significantly correlated with HER2 status determined by FISH, there are important differences in clinical samples that could have an impact on patient entry to clinical trials of HER2-targeted therapies. Recent data indicate that FISH is currently the most accurate method of determining HER2 status, and is therefore the preferable method for selecting patients for trastuzumab therapy.

Acknowledgements

We wish to acknowledge Jian-Yuan Zhou, Wen-Hsiang Wen, Brandon Li and Ivonne Villalobos at USC and Tajuana Everett, Sonja Faulk, and Lori Johnson at LabCorp for technical assistance.

Grant support: Michael F. Press received grant support from the National Cancer Institute (CA48780).

Presented in part at the 36th Annual Meeting of the American Society of Clinical Oncology, New Orleans, LA, May 20–23, 2000.

References

1. Akiyama T, Sudo C, Ogawara H, Toyoshima K, Yamamoto T: The product of the human c-erbB-2 gene: a 185-kilodalton glycoprotein with tyrosine kinase activity. Science 232: 1644–1646, 1986
2. Coussens L, Yang-Feng TL, Liao YC, Chen E, Gray A, McGrath J, Seeburg PH, Libermann TA, Schlessinger J, Francke U, et al.: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. Science 230: 1132–1139, 1985
3. Stern DF, Heffernan PA, Weinberg RA: p185, a product of the neu proto-oncogene, is a receptorlike protein associated with tyrosine kinase activity. Mol Cell Biol 6: 1729–1740, 1986
4. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL: Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235: 177–182, 1987
5. Pauletti G, Dandekar S, Rong H: A large-scale FISH study of HER-2/neu gene amplification in breast cancer. Proc Am Assoc Cancer Res 39: 345 (Abstract 2352), 1998
6. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A, et al.: Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science 244: 707–712, 1989
7. Kallioniemi OP, Kallioniemi A, Kurisu W, Thor A, Chen LC, Smith HS, Waldman FM, Pinkel D, Gray JW: ERBB2 amplification in breast cancer analyzed by fluorescence in situ hybridization. Proc Natl Acad Sci USA 89: 5321–5325, 1992
8. Gusterson BA, Gelber RD, Goldhirsch A, Price KN, Save-Soderborgh J, Anbazhagan R, Styles J, Rudenstam CM, Golouh R, Reed R, et al.: Prognostic importance of c-erbB-2 expression in breast cancer. International (Ludwig) Breast Cancer Study Group. J Clin Oncol 10: 1049–1056, 1992
9. Andrulis IL, Bull SB, Blackstein ME, Sutherland D, Mak C, Sidlofsky S, Pritzker KP, Hartwick RW, Hanna W, Lickley L, Wilkinson R, Qizilbash A, Ambus U, Lipa M, Weizel H, Katz A, Baida M, Mariz S, Stoik G, Dacamara P, Strongitharm D, Geddie W, McCready D: neu/erbB-2 amplification identifies a poor-prognosis group of women with node-negative breast cancer. Toronto Breast Cancer Study Group. J Clin Oncol 16: 1340–1349, 1998
10. Press MF, Pike MC, Chazin VR, Hung G, Udove JA, Markowicz M, Danyluk J, Godolphin W, Sliwkowski M, Akita R, et al.: Her-2/neu expression in node-negative breast cancer: direct tissue quantitation by computerized image analysis and association of overexpression with increased risk of recurrent disease. Cancer Res 53: 4960–4970, 1993
11. Press MF, Bernstein L, Thomas PA, Meisner LF, Zhou JY, Ma Y, Hung G, Robinson RA, Harris C, El-Naggar A, Slamon DJ, Phillips RN, Ross JS, Wolman SR, Flom KJ: HER-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas. J Clin Oncol 15: 2894–2904, 1997
12. Chazin VR, Kaleko M, Miller AD, Slamon DJ: Transformation mediated by the human HER-2 gene independent of the epidermal growth factor receptor. Oncogene 7: 1859–1866, 1992
13. Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL, Rowland AM, Kotts C, Carver ME, Shepard HM: Human-

ization of an anti-p185HER2 antibody for human cancer therapy. Proc Natl Acad Sci USA 89: 4285–4289, 1992
14. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, Wolter JM, Paton V, Shak S, Lieberman G, Slamon DJ: Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 17: 2639–2648, 1999
15. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 344: 783–792, 2001
16. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M, Shak S, Stewart SJ, Press M: Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 20: 719–726, 2002
17. Pauletti G, Godolphin W, Press MF, Slamon DJ: Detection and quantitation of HER-2/neu gene amplification in human breast cancer archival material using fluorescence in situ hybridization. Oncogene 13: 63–72, 1996
18. Pauletti G, Dandekar S, Rong H, Ramos L, Peng H, Seshadri R, Slamon DJ: Assessment of methods for tissue-based detection of the HER-2/neu alteration in human breast cancer: a direct comparison of fluorescence in situ hybridization and immunohistochemistry. J Clin Oncol 18: 3651–3664, 2000
19. Press MF, Hung G, Godolphin W, Slamon DJ: Sensitivity of HER-2/neu antibodies in archival tissue samples: potential source of error in immunohistochemical studies of oncogene expression. Cancer Res 54: 2771–2777, 1994
20. Tubbs RR, Pettay JD, Roche PC, Stoler MH, Jenkins RB, Grogan TM: Discrepancies in clinical laboratory testing of eligibility for trastuzumab therapy: apparent immunohistochemical false-positives do not get the message. J Clin Oncol 19: 2714–2721, 2001
21. Lebeau A, Deimling D, Kaltz C, Sendelhofert A, Iff A, Luthardt B, Untch M, Lohrs U: Her-2/neu analysis in archival tissue samples of human breast cancer: comparison of immunohistochemistry and fluorescence in situ hybridization. J Clin Oncol 19: 354–363, 2001
22. Jimenez RE, Wallis T, Tabasczka P, Visscher DW: Determination of Her-2/Neu status in breast carcinoma: comparative analysis of immunohistochemistry and fluorescent in situ hybridization. Mod Pathol 13: 37–45, 2000
23. Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P, Richmond J, Donald B, Ferrier R, Cooke TG: Evaluating HER2 amplification and overexpression in breast cancer. J Pathol 195: 422–428, 2001
24. Diaz NM: Laboratory testing for HER2/neu in breast carcinoma: an evolving strategy to predict response to targeted therapy. Cancer Control 8: 415–418, 2001
25. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ: Comparison of fluorescence in situ hybridization and immunohistochemistry for the evaluation of HER-2/neu in breast cancer. J Clin Oncol 17: 1974–1982, 1999
26. Kakar S, Puangsuvan N, Stevens JM, Serenas R, Mangan G, Sahai S, Mihalov ML: HER-2/neu assessment in breast cancer by immunohistochemistry and fluorescence in situ hybridization: comparison of results and correlation with survival. Mol Diagn 5: 199–207, 2000
27. Ridolfi RL, Jamehdor MR, Arber JM: HER-2/neu testing in breast carcinoma: a combined immunohistochemical and fluorescence in situ hybridization approach. Mod Pathol 13: 866–873, 2000
28. Perez EA, Roche PC, Jenkins RB, Reynolds CA, Halling KC, Ingle JN, Wold LE: HER2 testing in patients with breast cancer: poor correlation between weak positivity by immunohistochemistry and gene amplification by fluorescence in situ hybridization. Mayo Clin Proc 77: 148–154, 2002
29. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, Dupuis B, Parker RL: HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. Mod Pathol 14: 1079–1086, 2001
30. Hoang MP, Sahin AA, Ordonez NG, Sneige N: HER-2/neu gene amplification compared with HER-2/neu protein overexpression and interobserver reproducibility in invasive breast carcinoma. Am J Clin Pathol 113: 852–859, 2000
31. Seidman AD, Fornier MN, Esteva FJ, Tan L, Kaptain S, Bach A, Panageas KS, Arroyo C, Valero V, Currie V, Gilewski T, Theodoulou M, Moynahan ME, Moasser M, Sklarin N, Dickler M, D'Andrea G, Cristofanilli M, Rivera E, Hortobagyi GN, Norton L, Hudis CA: Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. J Clin Oncol 19: 2587–2595, 2001
32. Persons DL, Borelli KA, Hsu PH: Quantitation of HER-2/neu and c-myc gene amplification in breast carcinoma using fluorescence in situ hybridization. Mod Pathol 10: 720–727, 1997
33. Tsuda H, Akiyama F, Terasaki H, Hasegawa T, Kurosumi M, Shimadzu M, Yamamori S, Sakamoto G: Detection of HER-2/neu (c-erb B-2) DNA amplification in primary breast carcinoma. Interobserver reproducibility and correlation with immunohistochemical HER-2 overexpression. Cancer 92: 2965–2974, 2001
34. Birner P, Oberhuber G, Stani J, Reithofer C, Samonigg H, Hausmaninger H, Kubista E, Kwasny W, Kandioler-Eckersberger D, Gnant M, Jakesz R: Evaluation of the United States Food and Drug Administration-approved scoring and test system of HER-2 protein expression in breast cancer. Clin Cancer Res 7: 1669–1675, 2001
35. Press MF, Slamon DJ, Flom KJ, Park J, Zhou JY, Bernstein L: Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. J Clin Oncol 20: 3095–3105, 2002
36. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, Brown A, Yothers G, Anderson S, Smith R, Wickerham DL, Wolmark N: Real-world performance of HER2 testing – National Surgical Adjuvant Breast and Bowel Project experience. J Natl Cancer Inst 94: 852–854, 2002
37. Fleiss J: Statistical Methods of Rates and Proportions, 2nd edn New York: John Wiley & Sons, Inc, 1981, p. 217
38. Mass RD, Press MF, Anderson S: Evaluation of clinical outcomes according to HER2 detection by fluorescence *in situ* hybridization in women with metastatic breast cancer treated with trastuzumab. Clin Breast Cancer, 2005 (in press)
39. Ross JS, Fletcher JA: The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. Stem Cells 16: 413–428, 1998
40. Battifora H, Kopinski M: The influence of protease digestion and duration of fixation on the immunostaining of keratins. A comparison of formalin and ethanol fixation. J Histochem Cytochem 34: 1095–1100, 1986
41. Penault-Llorca F, Adelaide J, Houvenaeghel G, Hassoun J, Birnbaum D, Jacquemier J: Optimization of immunohistochemical detection of ERBB2 in human breast cancer: impact of fixation. J Pathol 173: 65–75, 1994
42. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ: Specificity of HercepTest in determining HER-2/neu status of breast cancers using the United States Food and Drug Administration-approved scoring system. J Clin Oncol 17: 1983–1987, 1999
43. Owens MA, Horten BC, Da Silva MM: HER2 amplification ratios by fluorescence in situ hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues. Clin Breast Cancer 5: 63–69, 2004
44. Press MF SG, Bernstein L, Mirlacher M, Villalobos I, Zhou J, Riva A, Nabholtz JM, Slamon DJ: Diagnostic evaluation of HER-2/neu as a molecular therapeutic target: local testing versus centralized FISH testing. Lab Invest 83: 43a, 2003

45. Niehans GA, Singleton TP, Dykoski D, Kiang DT: Stability of HER-2/neu expression over time and at multiple metastatic sites. J Natl Cancer Inst 85: 1230–1235, 1993
46. Vincent-Salomon A, Jouve M, Genin P, Freneaux P, Sigal-Zafrani B, Caly M, Beuzeboc P, Pouillart P, Sastre-Garau X: HER2 status in patients with breast carcinoma is not modified selectively by preoperative chemotherapy and is stable during the metastatic process. Cancer 94: 2169–2173, 2002
47. Simon R, Nocito A, Hubscher T, Bucher C, Torhorst J, Schraml P, Bubendorf L, Mihatsch MM, Moch H, Wilber K, Schotzau A, Kononen J, Sauter G: Patterns of her-2/neu amplification and overexpression in primary and metastatic breast cancer. J Natl Cancer Inst 93: 1141–1146, 2001
48. Masood S, Bui MM: Assessment of Her-2/neu overexpression in primary breast cancers and their metastatic lesions: an immunohistochemical study. Ann Clin Lab Sci 30: 259–265, 2000
49. Paik S, Tan-Chiu E, Bryant J, Romond E, Brown A, Mull J, Hiller W, Finnigan M, Wolmark N: Successful quality assurance program for HER2 testing in the NSABP trial for Herceptin. Breast Cancer Res Treat 76: S31, 2002
50. Bartlett J, Mallon E, Cooke T: The clinical evaluation of HER-2 status: which test to use? J Pathol 199: 411–417, 2003
51. CAP: Clinical laboratory assays for HER-2/neu amplification and overexpression: quality assurance, standardization, and proficiency testing. Arch Pathology Lab Med 126: 803–808, 2002
52. van Diest PJ vDP, Henzen-Logmans SC, Berns E, van der Burg ME, Green J, Vergote I: A scoring system for immunohistochemical staining: consensus report of the task force for basic research of the EORTC-GCCG. European Organization for Research and Treatment of Cancer-Gynaecological Cancer Cooperative Group. J Clin Pathol 50: 801–804, 1997
53. Going JJ, Mallon L, Reeves JR, Watters AD, Richmond J, Donald B, Ferrier R, Cooke T, Bartlett J: Inter-observer agreement in assessing C-ERBB-2 status in breast cancer: immunhistochemistry and FISH. J Pathol 190 (Suppl): 19A, 2000

Address for offprints and correspondence: Noël Dybdal, Genentech Inc., One DNA Way, MS 86, South San Francisco, CA 94080, USA; *Tel.:* +1-650-225-1425; *Fax:* +1-650-225-2797; *E-mail:* nod@gene.com