Justification of the Change From 10% to 30% for the Immunohistochemical HER2 Scoring Criterion in Breast Cancer

Yan-Hui Liu, MD,1* Fang-Ping Xu, MD, PhD,1* Jian-Yu Rao, MD,2 Heng-Guo Zhuang, MD,1 Xin-Lan Luo,1 Li Li, MS,1 Dong-Lan Luo, MS,1 Fen Zhang,1 and Jie Xu1

Key Words: HER2; Breast cancer; Immunohistochemistry; Fluorescence in situ hybridization

DOI: 10.1309/AJCPG3LCVKPNF J6R

Upon completion of this activity you will be able to:

• outline the protocol for the HER2 fluorescence in situ hybridization (FISH) test.
• describe the difference between Food and Drug Administration and American Society of Clinical Oncology/College of American Pathologists guidelines for the HER2 immunohistochemical scoring system for breast carcinomas.
• address interpretation of results using different criteria and cite factors that may lead to discordant results between immunohistochemical scoring and FISH for HER2 in breast carcinomas.

Abstract

We compared the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) immunohistochemical scoring criterion (30%) for determining HER2 status and the Food and Drug Administration criterion (10%) with fluorescence in situ hybridization (FISH), the HER2 gene amplification method in 328 cases of breast cancer. Of 294 tumor samples successfully analyzed simultaneously by FISH and immunohistochemically, 178 of 196 cases scored 3+ using the 10% and the 30% criteria. Using FISH as the reference, the number of false-positives was reduced from 24 to 9 after application of the 30% criterion. The specificity of immunohistochemical analysis was higher with the 30% (92.0%) vs the 10% (78.8%) criterion. The κ coefficient between FISH and immunohistochemical analysis was increased to 0.850 (almost perfect agreement; P < .001) after application of the 30% criterion vs 0.757 (substantial agreement) for the 10% criterion; the false-positive rate decreased to 5.1% from 12.2%. The χ² test showed that immunohistochemical analysis had significantly higher accuracy with the 30% (94.9%) vs the 10% (87.8%; P = .014) criterion. Our results from a large series of Chinese patients with breast cancer support that the ASCO/CAP 30% criterion may offer better results for assessing HER2 status.

It is well established that assessment of HER2 status is important in the management of breast cancer because HER2 status has prognostic implications, has predictive value for response to certain therapies, and is crucial for selecting candidates for trastuzumab therapy, an expensive and potentially toxic treatment.4-7 Immunohistochemical analysis and fluorescence in situ hybridization (FISH) are used for this purpose, but FISH is more expensive and requires specialized equipment.8-11 In the original standardized immunohistochemical testing algorithm for HER2 recommended by the US Food and Drug Administration (FDA), strong and complete membrane staining in more than 10% of invasive tumor cells was scored as 3+ and defined as HER2+.12 However, many recent reports showed that concordance between immunohistochemical and FISH testing was far from perfect and that many false-positive findings might be produced by using the FDA scoring system.8,13-15 Recently, the American Society of Clinical Oncology and College of American Pathologists (ASCO/CAP) guidelines have changed the criterion for scoring 3+ by immunohistochemical evaluation of HER2 status from 10% to 30%, which seemed to decrease the incidence of false-positive 3+ cases.16 However, a recent study showed that the 30% scheme was still too conservative.17 Beyond the 2 studies,16,17 to the best of our knowledge, there seems to be little published scientific justification evaluating the specificity and accuracy of the new immunohistochemical scoring guideline (30% criterion) for HER2.

In this study, to determine whether the new ASCO/CAP immunohistochemical scoring criterion for determining
HER2 status is justified, we used FISH as the assumed reference standard and sought to evaluate the correlation between immunohistochemical analysis and FISH in a series of 294 breast cancer cases, using the 2 sets of criteria for immunohistochemical analysis: the original criteria recommended by the FDA and the more recent ASCO/CAP guidelines.

Materials and Methods

Tumor Specimens

Invasive ductal breast carcinomas from 328 patients were used for the study. The cases were obtained from the archives of the Department of Pathology and Laboratory Medicine, Guangdong General Hospital, Guangzhou, China, between June 2005 and February 2008. None of the patients received adjuvant therapy before surgery. For the purposes of validation and assessment of specificity and accuracy of the new immunohistochemical scoring criterion, we intentionally included all cases with an immunohistochemical HER2 score of 3+ originally (n = 196) within the study period, whereas the remaining cases (n = 132) were randomly selected from a total of approximately 500 cases that had original immunohistochemical scores of 0 to 2+ within the same period. The specimens were obtained from paraffin blocks of 328 primary breast cancers, including 182 lymph node–negative and 146 lymph node–positive samples. Each specimen was fixed in neutral-buffered formalin and embedded in paraffin. The approval for this study was granted by the Guangdong General Hospital Medical Ethics Committee.

Tissue Microarray

Tissue microarrays (TMAs) containing 328 breast cancers were constructed from cases with adequate archival paraffin-embedded tissues. H&E-stained slides from each tumor block were reviewed and used as a guide to select areas representing potential heterogeneity of tumor cells from each case, which were sampled using a tissue arraying instrument (Beecher Instruments, Silver Spring, MD) to remove a 0.6-mm-diameter cylinder of tissue. The cylinder was then reembedded into a predetermined position in a recipient paraffin block. In these TMAs, for each of the 328 cases, 5 duplicate spots of primary tumor samples were sampled, yielding a total of 1,640 spots. Multiple sections (3-4 µm thick) were then cut from the TMA block and mounted on microscope slides.

Immunohistochemical Analysis

Tumor samples were assessed for HER2 status with a standardized in-house immunohistochemical protocol performed by certified laboratory staff using a DAKO autostaining kit (DAKO, Carpinteria, CA). Detailed methods are published elsewhere. Briefly, TMA sections (3 µm) were deparaffinized and rehydrated through graded alcohols. Endogenous peroxidase activity was blocked with 0.3% hydrogen peroxide for 10 minutes, and antigens were retrieved for 20 minutes in 10 mmol/L citrate buffer, pH 6.0, at 98°C. Nonspecific binding was blocked with 10% normal goat serum for 20 minutes. The TMA slides were incubated with the rabbit antihuman HER2 polyclonal antibody A0485 (1:500 dilution; DAKO) for 65 minutes at 25°C. Then the slides were incubated with a ChemMate DAKO Envision system (K5007) for 30 minutes at 25°C and subsequently reacted with 3,3′-diaminobenzidine as a chromogen substrate. The nucleus was counterstained using Mayer hematoxylin. The negative control experiment was performed by replacing the primary antibody with a normal murine IgG. Known immunostaining slides were used as positive control samples.

Staining intensity was graded according to the original FDA12 and the new ASCO/CAP criteria, respectively.16 In brief, according to the FDA guidelines, scoring was performed as follows: score 0 was given for no staining or staining in fewer than 10% of the invasive tumor cells, 1+ for weak and incomplete membrane staining in more than 10% of the invasive tumor cells, 2+ for tumors with weak to moderate complete membrane staining in more than 10% of the cells, and 3+ for strong and complete membrane staining in more than 10% of the invasive tumor cells. According to the new ASCO/CAP guidelines, scores 0, 1+, and 2+ are same as the FDA system, whereas a score of 3+ is given for strong and complete membrane staining in more than 30% of the invasive tumor cells. Interpretations using the FDA and ASCO/CAP guidelines were done by 2 pathologists (Y.-H.L. and H.-G.Z.) independently, both of whom were completely blinded to the FISH results. Discrepant cases were concurrently reviewed by these 2 pathologists using a multiheaded microscope, where a consensus score was reached.

Fluorescence In Situ Hybridization

The FDA-approved PathVysion HER2 DNA probe kit (Abbott Molecular, Des Plaines, IL) was used in accordance with the manufacturer’s recommended protocols but with some minor modifications. Briefly, the deparaffinized TMA section was treated in a pretreatment solution at 85°C for 20 minutes followed by protease digestion at 37°C for 8 minutes. Next, 20 µL of hybridization solution containing directly labeled probes, both SpectrumGreen for the chromosome 17 centromere (CEP) and SpectrumOrange for the HER2 gene locus, was applied, and the probe–target tissue was codecked as follows: score 0 was given for no staining or staining in fewer than 10% of the invasive tumor cells, 1+ for weak and incomplete membrane staining in more than 10% of the invasive tumor cells, 2+ for tumors with weak to moderate complete membrane staining in more than 10% of the cells, and 3+ for strong and complete membrane staining in more than 10% of the invasive tumor cells. According to the new ASCO/CAP guidelines, scores 0, 1+, and 2+ are same as the FDA system, whereas a score of 3+ is given for strong and complete membrane staining in more than 30% of the invasive tumor cells. Interpretations using the FDA and ASCO/CAP guidelines were done by 2 pathologists (Y.-H.L. and H.-G.Z.) independently, both of whom were completely blinded to the FISH results. Discrepant cases were concurrently reviewed by these 2 pathologists using a multiheaded microscope, where a consensus score was reached.
(Zeiss, Göttingen, Germany) equipped with a triple bandpass filter. Lymphocytes and normal breast tissue served as normal control samples. A known paraffin-embedded primary breast cancer specimen with HER2 amplification was selected as a positive control sample.

FISH results were recorded in a data report form, listing the number of red signals for the probe (HER2) and the number of green signals for CEP17 for each nucleus investigated, with a total of 20 nuclei for each sample. The presence of CEP17 signals in at least 75% of the cancer cell nuclei was confirmed. The area of invasive carcinoma for analysis was marked on the reverse surface with a diamond pen, as determined by review before and after H&E staining. Signals were counted in nuclei with identifiable boundaries. Optimally, only signals distinctly separated from each other were included, but in cases with high levels of amplification, the signals formed clusters and the number had to be estimated. The FISH analyses for HER2 were performed independently and without knowledge of the immunohistochemical result. The HER2/CEP17 ratios were calculated. As stated in the manufacturer’s guidelines (or the new ASCO/CAP guidelines), a sample with an HER2/CEP17 ratio greater than 2.2 was considered amplified, a ratio less than 1.8 was considered nonamplified, and a ratio between 1.8 and 2.2 was considered equivocal.

### Statistical Analyses

Statistical analysis was performed with the SPSS software (SPSS Standard, version 10.0, SPSS, Chicago, IL). PathVysion was used as the presumed reference standard against which the different criteria for immunohistochemical analysis were compared, and agreement was quantified with pairwise κ statistics. The McNemar test was used to assess overvaluation or undervaluation of HER2 status. Categorical data were compared by using the χ² or Fisher exact test. Statistical significance was assumed if the P value was less than .05.

### Results

HER2 FISH analysis was successfully performed in 294 (89.6%) of 328 cases. Samples without a FISH signal and samples with weak target signals and strong background signals were the reasons for most of the noninformative cases. A small portion of noninformative cases were due to problems related to the TMA technology, including lost samples, unrepresentative samples, and samples with too few tumor cells; such samples were not used in data compilation. Of the 294 tumors, 181 (61.6%) tumors were identified as amplified, 4 (1.4%) as equivocal, and 109 (37.1%) as nonamplified.

Of the 294 tumor samples successfully analyzed simultaneously by FISH and immunohistochemical analysis, 32 were scored as 0, 27 as 1+, 39 as 2+, and 196 as 3+ using the original FDA 10% criterion. The cases amplified by FISH according to the immunohistochemical score were as follows: 3 (9%) amplified with an immunohistochemical score of 0; 2 (7%) amplified with an immunohistochemical score of 1+; 4 (10%) amplified with an immunohistochemical score of 2+; and 172 (87.8%) amplified with an immunohistochemical score of 3+. In addition, 1 case scored 2+ and 3 cases scored 3+ were identified as equivocal by FISH. After application of the ASCO/CAP 30% criterion, 18 cases received different scores compared with the 10% criterion Table II, which were all reanalyzed on routine tissue sections. The 18 cases were scored 3+ with the 10% criterion but 2+ with the 30% criterion. By using the 30% criterion, the immunohistochemical scores for the 294 cases were as follows: 178 cases, 3+; 57 cases, 2+; 27 cases, 1+; and 32 cases, 0; the latter 2 groups were the same as with the 10% criterion. The FISH assay showed that 7 of 57 cases with an immunohistochemical score of 2+ were amplified, and 2 cases were equivocal; 169 of 178 cases scored immunohistochemically as 3+ were amplified, and 2 cases also were equivocal.

With PathVysion FISH as the reference standard, the concordance between FISH and the 2 sets of immunohistochemical results are summarized in Table II and Table III; the equivocal cases were considered negative by FISH. Concordance between FISH and immunohistochemical results with the 10% criterion was 88.8%. The κ coefficient was calculated as 0.757 (P < .001). This coefficient was increased to 0.850 (almost perfect agreement; P < .001) after application of the 30% criterion, with an overall concordance of 92.9%. Interobserver

Table II

<table>
<thead>
<tr>
<th>Case No.</th>
<th>10% Criterion</th>
<th>30% Criterion</th>
<th>FISH Ratio</th>
<th>Chromosome 17 Polysomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3</td>
<td>2</td>
<td>1.9</td>
<td>Absent</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>2</td>
<td>1.3</td>
<td>Absent</td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>2</td>
<td>1.1</td>
<td>Present</td>
</tr>
<tr>
<td>62</td>
<td>3</td>
<td>2</td>
<td>2.5</td>
<td>Absent</td>
</tr>
<tr>
<td>71</td>
<td>3</td>
<td>2</td>
<td>1.3</td>
<td>Absent</td>
</tr>
<tr>
<td>78</td>
<td>3</td>
<td>2</td>
<td>1.2</td>
<td>Absent</td>
</tr>
<tr>
<td>81</td>
<td>3</td>
<td>2</td>
<td>1.5</td>
<td>Present</td>
</tr>
<tr>
<td>146</td>
<td>3</td>
<td>2</td>
<td>1.3</td>
<td>Absent</td>
</tr>
<tr>
<td>195</td>
<td>3</td>
<td>2</td>
<td>1.2</td>
<td>Absent</td>
</tr>
<tr>
<td>216</td>
<td>3</td>
<td>2</td>
<td>1.6</td>
<td>Present</td>
</tr>
<tr>
<td>222</td>
<td>3</td>
<td>2</td>
<td>1.1</td>
<td>Absent</td>
</tr>
<tr>
<td>240</td>
<td>3</td>
<td>2</td>
<td>2.8</td>
<td>Absent</td>
</tr>
<tr>
<td>242</td>
<td>3</td>
<td>2</td>
<td>1.4</td>
<td>Absent</td>
</tr>
<tr>
<td>244</td>
<td>3</td>
<td>2</td>
<td>1.0</td>
<td>Absent</td>
</tr>
<tr>
<td>268</td>
<td>3</td>
<td>2</td>
<td>1.6</td>
<td>Present</td>
</tr>
<tr>
<td>274</td>
<td>3</td>
<td>2</td>
<td>1.3</td>
<td>Absent</td>
</tr>
<tr>
<td>290</td>
<td>3</td>
<td>2</td>
<td>1.4</td>
<td>Absent</td>
</tr>
<tr>
<td>291</td>
<td>3</td>
<td>2</td>
<td>9.4</td>
<td>Absent</td>
</tr>
</tbody>
</table>

FISH, fluorescence in situ hybridization.

* The numbers are immunohistochemical scores.

© American Society for Clinical Pathology
concordance for cases scored immunohistochemically as 3+ was good for the FDA (κ = 0.815) and ASCO/CAP guidelines (κ = 0.843). The sensitivity for immunohistochemical determination of HER2 status using the 10% criterion was 95.0%, and the specificity was 78.8%. Immunohistochemical analysis with the 30% criterion produced a sensitivity decreasing to 93.4% and a specificity increasing to 92.0%. The false-positive rate was 12.2%, and the false-negative rate was 9.2% for immunohistochemical analysis with the 10% criterion, whereas the false-positive rate was 5.1% and the false-negative rate 10.3% when using the 30% criterion.

The McNemar test indicated that there was no significant difference between FISH and immunohistochemical analysis using the 30% criterion (P = .664), but there was a significant difference between FISH and immunohistochemical analysis with the 10% criterion (P = .014). Furthermore, the χ² test showed that the results for HER2 status for immunohistochemical scores of 3+ using the 2 criteria confirmed by FISH were significantly different: immunohistochemical analysis with the 30% criterion (94.9%) had a significantly higher accuracy than that with the 10% criterion (87.8%; P = .014).

When equivocal FISH cases were analyzed as a separate group (negative, equivocal, positive) rather than considered as negative, we found that with the 10% cutoff, 26 (8.8%) of 294 cases showed discordance between immunohistochemical analysis and FISH (5 negative by immunohistochemical analysis and positive by FISH; 21 positive by immunohistochemical analysis and negative by FISH). With the 30% cutoff, only 12 (4.1%) of 294 cases were discordant (5 negative by immunohistochemical analysis and positive by FISH; 7 positive by immunohistochemical analysis and negative by FISH). Thus, changing to the 30% cutoff reduced the discordant rate by half.

**Discussion**

Accurate assignment of the HER2 status of invasive breast cancers in formalin-fixed, paraffin-embedded samples has been questioned, and numerous studies have addressed discordant results between immunohistochemical analysis and FISH. Since our laboratory has active quality assurance and quality control programs for routine HER2 FISH and immunohistochemical assays, we found that the interpretation of results using different criteria might also be a key factor.

In the present study, we assessed the specificity and accuracy of the new ASCO/CAP immunohistochemical scoring system (30% criterion) by comparison with FISH in 5 separate TMA cores from each tumor. It should be indicated that while FISH results were used as the assumed reference standard for determination of HER2 status in the current study, such a notion is by no means universally accepted. The use of TMA is similar, in some ways, to examination of 4 to 5 mid power (×20) microscopic fields. In our cohorts of breast cancer, after using the 30% criterion, the discordance rate was reduced from 12.2% to 5.1% for the cases 3+ immunohistochemically, and the specificity of immunohistochemical analysis was increased to 92.0% from 78.8%. Moreover, there was a marked improvement in the total concordance rate between FISH and immunohistochemical analysis after using the 30% criterion (κ = 0.850; almost perfect agreement) compared with using the 10% criterion (κ = 0.757; substantial agreement). In addition, the McNemar test indicated that there was no significant difference between FISH and immunohistochemical results with the 30% criterion, but there was a significant difference between FISH and immunohistochemical results with the 10% criterion. This finding indicated that immunohistochemical analysis with the 30% criterion could more accurately estimate HER2 status and reduce the false-positive rate than use of the 10% criterion.

During the preparation of the present report, 2 independent groups also demonstrated that there was a greater concordance coefficient between FISH and immunohistochemical analysis with the 30% criterion in 100 and 98 cases of invasive breast cancers, respectively. Together with those studies, we conclude that immunohistochemical analysis with

---

**Table 2**

<table>
<thead>
<tr>
<th>Immunohistochemical Result</th>
<th>FISH+</th>
<th>FISH−</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (3+)</td>
<td>172</td>
<td>24</td>
<td>196</td>
</tr>
<tr>
<td>Negative (0, 1+, 2+)</td>
<td>9</td>
<td>89</td>
<td>98</td>
</tr>
<tr>
<td>Total</td>
<td>181</td>
<td>113</td>
<td>294</td>
</tr>
</tbody>
</table>

FISH, fluorescence in situ hybridization.

* Concordance data, given as estimate (95% confidence interval), are as follows: κ coefficient, 0.757 (0.679-0.835); sensitivity, 95.0% (91.9%-98.1%); specificity, 78.8% (71.4%-86.2%); and overall, 88.5% (85.3%-92.3%). Overall concordance = (True Positive + True Negative)/Total Cases. True positives are cases that are both immunohistochemically positive and FISH+; true negatives are cases that are both immunohistochemically negative and FISH−.

**Table 3**

<table>
<thead>
<tr>
<th>Immunohistochemical Result</th>
<th>FISH+</th>
<th>FISH−</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (3+)</td>
<td>169</td>
<td>9</td>
<td>178</td>
</tr>
<tr>
<td>Negative (0, 1+, 2+)</td>
<td>12</td>
<td>104</td>
<td>116</td>
</tr>
<tr>
<td>Total</td>
<td>181</td>
<td>113</td>
<td>294</td>
</tr>
</tbody>
</table>

FISH, fluorescence in situ hybridization.

* Concordance data, given as estimate (95% confidence interval), are as follows: coefficient, 0.850 (0.787-0.913); sensitivity, 93.4% (89.9%-96.9%); specificity, 92.0% (87.1%-96.9%); and overall, 92.9% (90.0%-95.8%). Overall concordance = (True Positive + True Negative)/Total Cases. True positives are cases that are both immunohistochemically positive and FISH+; true negatives are cases that are both immunohistochemically negative and FISH−.
the 30% criterion may offer better results for the assessment of HER2 status.

Further analysis of the 18 discordant cases revealed that chromosome 17 polysomy accounted for 4 cases, and false-positive immunohistochemical results with the 10% criterion, owing to rarity of tumor cells and intratumor heterogeneity, may be important reasons for some of the discrepancies.\(^{15-17,20}\) However, it is also important to note that there were 3 cases in which FISH demonstrated amplification and the immunohistochemical result was only 2+ using the 30% criterion. These cases may indeed be false-negatives with the 30% criterion (undervaluation). Alternative explanations include low-level amplification with FISH, which was observed in 1 case (ratio, 2.5). An additional case with an immunohistochemical score of 2+ was scored as equivocal by FISH. The borderline FISH result is known to be technically challenging.\(^{13-16}\) wherein interobserver variability may have a role. Another reason, though less likely, may be a false-positive finding by FISH.

In our study, HER2 overexpression (3+ immunohistochemically) in the absence of amplification was observed in 9 cases, even using the 30% criterion and repeated FISH analysis of the routine tissue sections from the tumors. In fact, this is not surprising considering that there may be many variables involved in the immunohistochemical and FISH methods.\(^{18-21}\) Chromosome 17 polysomy (2 cases), low-level amplification (2 cases), and false-negative FISH results may explain the differences.

By using FISH as the reference standard, we have shown that immunohistochemical analysis with the 30% criterion provides better specificity and accuracy for the detection of HER2 status in breast cancer than the previously used 10% criterion. In a large series of Chinese patients with breast cancer, our results provide evidence to support the new ASCO/CAP guidelines\(^6\) for the assessment of HER2 status. With this approach, however, some false-negative cases may be problematic, and it may be advisable to submit samples with 2+ immunohistochemical scores for FISH testing to avoid false-positive and false-negative immunohistochemical results.

References


