Effects of the Change in Cutoff Values for Human Epidermal Growth Factor Receptor 2 Status by Immunohistochemistry and Fluorescence In Situ Hybridization

A Study Comparing Conventional Brightfield Microscopy, Image Analysis-Assisted Microscopy, and Interobserver Variation

Roscoe Atkinson, MD; Jens Mollerup, PhD; Anne-Vibeke Lænkholm, MD; Mark Verardo, PhD; Debra Hawes, MD; Deborah Commins, MD, PhD; Birte Engvad, MD; Adrian Correa, MD, MBA; Charlotte Cort Ehlers, MD; Kirsten Vang Nielsen, PhD

Context.—New guidelines for HER2 testing have been introduced.

Objectives.—To evaluate the difference in HER2 assessment after introduction of new cutoff levels for both immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) and to compare interobserver agreement and time to score between image analysis and conventional microscopy.

Design.—Samples from 150 patients with breast cancer were scored by 7 pathologists using conventional microscopy, with a cutoff of both 10% and 30% IHC-stained cells, and using automated microscopy with image analysis. The IHC results were compared individually and to HER2 status as determined by FISH, using both the approved cutoff of 2.0 and the recently introduced cutoff of 2.2.

Results.—High concordance was found in IHC scoring among the 7 pathologists. The 30% cutoff led to slightly fewer positive IHC observations. Introduction of a FISH equivocal zone affected 4% of the FISH scores. If cutoff for FISH is kept at 2.0, no difference in patient selection is found between the 10% and the 30% IHC cutoff. Among the 150 breast cancer samples, the new 30% IHC and 2.2 FISH cutoff levels resulted in one case without a firm diagnosis because both IHC and FISH were equivocal. Automated microscopy and image analysis-assisted IHC led to significantly better interobserver agreement among the 7 pathologists, with an increase in mean scoring time of only about 30 seconds per slide.

Conclusions.—The change in cutoff levels led to a higher concordance between IHC and FISH, but fewer samples were classified as HER2 positive.

(arch pathol lab med. 2011;135:1010–1016)

The HER2 gene, with the official gene name ERBB2, is amplified1 and overexpressed2 in 15% to 25% of breast cancers.3 The highest frequency is found among patients eligible for chemotherapy,3 whereas frequencies as low as 7% have been reported among patients receiving endocrine treatment.4 The efficacy of HER2-targeting therapy was first demonstrated in metastatic breast cancer,5 and later, the adjuvant effect in early breast cancer was established.6,7 With the introduction of trastuzumab (Herceptin) in 1998, accurate assessment of HER2 status became essential for the clinical management of patients with breast cancer. The companion diagnostic test, the HercepTest (Dako Denmark A/S, Glostrup, Denmark), classifies the HER2 membrane staining into 4 categories: 3+ is a positive result, with strong, complete membrane staining; 2+ is weakly positive or equivocal, with weak to moderate, complete membrane staining; 1+ is a negative result, with faint incomplete membrane staining; and 0 is negative, with no membrane staining. Initially, treatment with trastuzumab was recommended in the United States to patients with both immunohistochemistry (IHC) scores of 2+ and 3+, but retrospective analyses have suggested that only patients with 3+ score by IHC, and/or gene amplification, benefited.2 The debate about the accuracy of HER2 assessment was initiated by reports on the lack of
HER2 amplification in most 2+ cases. When trastuzumab was later registered in the European Union, the recommendation was restricted to patients with an HER2 IHC score of 3+ or 2+ with confirmed HER2 gene amplification. Several methods can be used to determine HER2 overexpression or gene amplification but the 2 most relevant diagnostic methods currently in use are IHC and fluorescence in situ hybridization (FISH). There is still no consensus on which of these methods is the most predictive. Since 1998, HER2 testing has been based on the HercepTest guidelines approved by the US Food and Drug Administration (FDA). Other IHC tests have been approved by concordance to the HercepTest. The FISH scoring guidelines for selection of patients for trastuzumab treatment was later approved by the FDA using a HER2 gene to centromere 17 ratio (HER2/CEN-17) cutoff for HER2 gene amplification of 2.0. Almost 10 years after the introduction of HER2 testing, joint guidelines were introduced by the American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) for the laboratory evaluation of HER2 assessment. The new guidelines were introduced because of discordant testing results among local and central laboratories and to achieve a higher concordance between IHC and FISH results. These new guidelines recommended validation of the assays, use of standard operating procedures, use of proficiency testing and laboratory accreditation standards, and importantly, new cutoff values for considering a tumor HER2 positive.

The new ASCO/CAP guidelines redefine a HER2-positive result as 3+ staining by IHC of greater than 30% of invasive tumor cells, as compared with the 10% cutoff used in the clinical trials leading to approval of trastuzumab. The earlier FISH cutoff of 2.0 has been substituted with an equivocal zone of 1.8 to 2.2 that requires additional testing. However, it is not known how many patients the new guidelines will affect by providing either a different, or a double equivocal, diagnosis.

To investigate the effect of part of these new guidelines, and especially the new cutoff values (ie, 3+ at 30%), we initiated a study of 150 invasive breast cancer specimens, interpreted by the FDA-approved HER2 FISH pharmDx assay (Dako) and the HercepTest, and scored by 7 different pathologists. The results were compared with the ASCO/CAP criteria for HER2 FISH and to a 30% cutoff (30% cutoff for all scores: 1+, 2+, and 3+). Although we are not advocating the use of a general 30% cutoff, it was useful to investigate the 3+ cutoff at the 30% level. Additional information was gathered, including the (1) concordance between the HercepTest and the HER2 FISH pharmDx assays, (2) comparison of manual analysis versus image analysis-assisted scoring, and (3) comparison of the time taken for image analysis scoring versus manual slide scoring.

**MATERIALS AND METHODS**

**Tumor Specimens**

Routine, formalin-fixed, paraffin-embedded breast cancer specimens from 150 patients were included in this study. Full serial sections were cut at 4 μm to 5 μm. Breast cancer specimens were selected to get a distribution of HER2 protein expression levels close to that observed in women newly diagnosed with primary breast cancer.

**Fluorescence In Situ Hybridization**

The HER2 FISH pharmDx assay was used according to the manufacturer’s recommendations. Scoring was performed by an experienced technologist using an Olympus BX51 fluorescence microscope (Olympus Denmark A/S, Ballerup, Denmark), equipped with appropriate filters for 4′,6-diamidino-2-phenylindole (DAPI), fluorescein isothiocyanate (FITC), and Texas Red. All borderline cases were rescoring by a more-experienced technologist. Verification was done through random sample checks by an experienced pathologist on a separate microscope system.

**Immunohistochemistry**

Preparation and staining of sections were performed according to the package insert of the HercepTest. Pretreatment and staining of HercepTest slides were performed using PT Link and Autostainer Plus Link (Dako). One section of every specimen was stained with hematoxylin-eosin and was available for inspection during the scoring of FISH and HercepTest slides.

**Conventional and Image Analysis-Assisted Slide Scoring Using Different Guidelines**

Seven pathologists, representing different levels of experience in reading HercepTest slides, scored the HercepTest-stained breast cancer slides using conventional brightfield microscopy. Before scoring, all pathologists were instructed on the study setup and were trained using a training set. The HercepTest was scored according to the FDA-approved manufacturer guidelines, with a cutoff of 10% and a cutoff of 30% of the invasive cancer cells having positive membrane staining. Both cutoff levels were applied for all scores: 1+, 2+, and 3+. Pathologists were also allowed to comment on each slide, if applicable.

The HER2 slides were also scanned on an Automated Cellular Imaging System (ACIS III; Dako). These digitized HER2 slides were saved and preloaded onto ACIS systems or ACIS workstations. Pathologists were first given an introduction to the ACIS, including the use of a training set if they were unfamiliar with the ACIS. Following this, the slides were scored. Briefly, areas of interest were selected in the low-resolution images, and the high-resolution images were displayed in the viewer box. The display magnification could be adjusted to the desired magnification, and a hot-spots function could be used. The 40×-circle drawing tool was used to demarcate the area of interest on a high-resolution image, which was then automatically scored. The pathologists were instructed to always examine the image to verify the ACIS results. At least 6 regions per slide were to be scored in this manner. The pathologists were instructed to examine the ACIS overview; to place a checkmark in a box corresponding to less than 10%, greater than 10%, less than 30%, and/or greater than 30% membrane staining; and to make comments, if appropriate, for each slide on a report form. This also allowed the pathologist to put in a score different from the ACIS-generated fractional score if he or she did not agree with the ACIS fractional score. If a pathologist did not score a minimum of 6 regions, the results were not used in the analysis. If a pathologist put in a score different from the ACIS score, that pathologist’s score was used for the final analysis. Following a pathologist’s scoring, the ACIS region scores were imported to an Excel (Microsoft Corporation, Redmond, Washington) spreadsheet and used for data analysis. Average scores were generated in Excel. Scoring time per batch of 10 slides for manual microscopy (2 scorings) and scoring time with the ACIS III were tracked using a digital stopwatch and recorded.

**Conversion of HER2 IHC Scores and HER2/CEN-17 Ratios to HER2 Status**

Manual HercepTest and HER2 FISH pharmDx scores were converted to HER2 status as shown in Table 1. Additionally, the population with HER2-positive status was defined either as positive with the HercepTest (using a 10% cutoff according to the...
manufacturer guidelines or a 30% cutoff according to the ASCO/CAP guidelines for a 3+ score or as positive by FISH (following Dako guidelines; ie, a HER2/CEN-17 score ≥2.0). In addition, the population with HER2-negative status was defined as those samples that were negative either by having a negative HercepTest result (manufacturer guidelines or using a 30% cutoff; ie, a 0 or 1+ score by any of the pathologists) or by having a negative FISH result (Dako guidelines; ie, a HER2/CEN-17 score <2.0). Fractional scores from the ACIS III system were converted to ACIS scores according to the application guide. In short, the fractional scores were rounded to the nearest whole number (ie, ACIS III fractional score ≥2.5 represented 3+ result).

### Results

**Concordance Between the HercepTest and the HER2 FISH PharmDx Assays**

Samples from 150 patients with breast cancer were analyzed by 7 pathologists using manual HercepTest scoring and both a 10% cutoff and a 30% cutoff, which were compared with the FISH scoring according to manufacturer and ASCO/CAP guidelines. Each pathologist was able to score between 136 and 150 samples by manual IHC and slightly fewer by ACIS-assisted scoring (130–148). Of 1050 possible scores, 997 data points (95%) were available (Table 2), and 53 scores (5%) were removed from the data sets and analysis. Table 2 shows the comparison between HER2 status obtained by manual IHC using the cutoffs from the 2 guidelines. The overall agreement is 94.4% (941 of 997). When excluding the equivocal or 2+ cases (n = 192) from these 997 scores, there was 100% agreement. The 30% cutoff caused more negative HER2 status scores, and fewer scores were positive or had equivocal status (Table 2).

The ASCO/CAP guidelines have introduced an equivocal range from 1.8 to 2.2 in HER2 FISH scores. According to the guidelines, cases in the range of 2.0 to 2.2, which are positive using the FDA-approved Dako guidelines, are equivocal using the ASCO/CAP guidelines. Introducing a 2.2 cutoff for positivity instead of 2.0 influenced 5 specimens (3%) that could potentially have their results changed. If FISH had been the entry test, those 5 samples would have been classified HER2 positive, with a ratio in the range of 2.0 to 2.2. When compared by IHC, only 3 discordant scores (0.03%) were identified, and they were not systematic because they represented 3 different samples scored by 3 different pathologists.

Using the criteria for “positive scores” and “negative scores” defined in Table 1 (ie, cutoff for a positive HER2 FISH result, 2.0), the number of IHC-positive-and-FISH-negative results and the number of IHC-negative-and-FISH-positive scores were investigated using the 2 different IHC cutoff values. Among the 262 positive scores, a high concordance was found between the 10% and the 30% cutoff (96.2% agreement; χ value, 0.92; Table 3). Applying a 30% cutoff, instead of a 10% cutoff, 5 observations changed from a positive to an equivocal HER2 status result. These observations would have been redirected to HER2 FISH and thereby assigned to the appropriate category of positive cases eligible for trastuzumab treatment. Table 3 also shows that a high concordance is found among the 803 negative scores (93.3% agreement).
Table 3. Comparison of Positive and Negative Scores in Human Epidermal Growth Factor Receptor 2 (HER2) Status for Manual Microscopy Using 10% and 30% Cutoffs

<table>
<thead>
<tr>
<th>Manual IHC Results, 30% Cutoff</th>
<th>HER2 Status, 10% Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (0/1+)</td>
<td>49</td>
</tr>
<tr>
<td>Positive (3+)</td>
<td>0</td>
</tr>
<tr>
<td>Equivocal (2+)</td>
<td>0</td>
</tr>
<tr>
<td>Total positive scores</td>
<td>49</td>
</tr>
<tr>
<td>Negative (0/1+)</td>
<td>630</td>
</tr>
<tr>
<td>Positive (3+)</td>
<td>0</td>
</tr>
<tr>
<td>Equivocal (2+)</td>
<td>1</td>
</tr>
<tr>
<td>Total negative scores</td>
<td>631</td>
</tr>
</tbody>
</table>

Abbreviation: IHC, immunohistochemistry.

agreement; κ value, 0.78). The main difference is that 50 assessments scored as equivocal results with a 10% cutoff are scored as negative results with a 30% cutoff.

The overall agreement between manual IHC and FISH can be seen in Table 4. Cases with an equivocal IHC (2+) are not included, and in the ASCO/CAP guidelines, equivocal FISH scores are not included. Apparently a slightly better agreement (95.0% agreement; κ value, 0.85), using the 30% cutoff and 2.2 FISH ratio cutoff, is obtained by manual scoring compared with the agreement (92.1% agreement; κ value, 0.78) when applying the FDA-approved manufacturer guidelines of 10% stained cells and 2.0 FISH ratio cutoff. However, these results were caused by the exclusion of a few difficult cases that ranged between 1.6 and 2.2 using FISH (excluded according to ASCO/CAP guidelines) and by the exclusion of more 2+ cases when the cutoff for a 3+ result was increased to 30%.

Consensus IHC Scoring Compared With FISH

A consensus IHC score was defined as a concordant result from at least 4 of the pathologists, which lead to 140 (93%) and 138 (92%) consensus IHC scores for manual IHC with 10% and 30% cutoffs, respectively. One hundred thirty-one cases (87%) were evaluated with agreement among all 7 pathologists, whereas in only 6 (4%) of the evaluated cases did 1, 2, or 3 pathologists disagree. The remaining 13 cases (9%) did not have a consensus IHC score (<4 pathologists agreed). Those that lacked a consensus IHC score were primarily caused by most of the pathologists failing to provide a score for those cases. Table 5 shows the IHC HER2 status using the HER2 FISH status using the Dako guidelines. Table 6 shows the IHC HER2 status using the 30% cutoff levels, compared with the HER2 FISH status (using ASCO/CAP guidelines). Overall, 24.3% of the cases were amplified using the Dako/FDA cutoff of 2.0, whereas only 21.0% were amplified using the ASCO/CAP cutoff of 2.2. According to the Dako guidelines and 10% cutoff value, 28 patients would have been eligible for treatment with trastuzumab (ie, 3+ IHC and/or FISH amplified) if IHC was the entry analysis, and an additional 6 patients would have been eligible for treatment if FISH was chosen for the initial analysis (Table 5). A negative HER2 FISH status was given to 106 and 103 cases, respectively; both guidelines were associated with only one IHC-positive-FISH-negative assessment, with a FISH ratio of 1.06 (Tables 5 and 6). According to the ASCO/CAP FISH guidelines and the 30% cutoff, 26 patients would have been eligible for treatment with trastuzumab if IHC was the entry analysis, and an additional 3 patients would have been eligible for treatment if FISH was chosen as the entry analysis (Table 6). This influences the HER2 assessment of at least one case with a FISH ratio of 2.0 and classified as an IHC-positive result with the Dako guidelines, but an IHC- and FISH-equivocal result using the ASCO/CAP FISH guidelines and the 30% cutoff (Table 7).

Overall, 90 cases were negative and 24 cases were positive, when comparing the consensus IHC HER2 status

Arch Pathol Lab Med—Vol 135, August 2011
The IHC status is based on consensus of scores by 7 pathologists. Values are shown in Table 4. When using image analysis-assisted microscopy with a cutoff of 10%, overall agreement was very similar (Table 4) to the overall agreement obtained by manual microscopy. A paired t test showed no significant difference between conventional microscopy and image analysis-assisted microscopy in concordance between HER2 status (positive/negative) obtained using HercepTest (10% cutoff) and HER2 FISH (2.0 cutoff; P = .70) when 2+ equivocal cases were excluded. This paired t test analysis was performed by comparing the 2 sets of κ values obtained from the stratified comparisons of conventional microscopy versus HER2 FISH and image analysis-assisted microscopy versus HER2 FISH (the averages of the κ values are shown in Table 4).

Comparison of Scoring Time of HercepTests Between Automated Image Analysis and Conventional Microscopy

A difference in the scoring time for HercepTests was found among pathologists and between conventional microscopy and image analysis-assisted microscopy. The mean scoring times for the individual pathologists are plotted in the Figure. The mean (SD) scoring time (minute:seconds) was 1:01 (0:26) using conventional microscopy and 1:31 (0.38) using ACIS-assisted scoring. The range was from 0:17 to 2:22 for manual scoring and from 0:37 to 2:58 for image analysis-assisted microscopy. A 2-tailed paired t test of the mean scoring times revealed a significant difference (P < .001). For one pathologist, a shorter scoring time using image analysis was observed, and for another, no difference in time to score was observed.

**Comment**

The present study revealed a good concordance in the IHC scores among the 7 pathologists with both the 10% and a 30% cutoff values, and a consensus HER2 IHC assessment could be given to 140 (93%) and 138 (92%) of

<table>
<thead>
<tr>
<th>FISH Ratio</th>
<th>IHC 10%*</th>
<th>IHC 30%*</th>
<th>FISH Cutoff 2.0</th>
<th>FISH Cutoff 2.2</th>
<th>Effect of Change in FISH Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>Pos</td>
<td>Equ (2+) Pos</td>
<td>Equ</td>
<td>No firm HER2 status because of double-equivocal diagnoses</td>
<td></td>
</tr>
<tr>
<td>2.03</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Equ</td>
<td>The 4 negative results by IHC (both cutoffs) are changed from FISH-positive to FISH-equivocal results</td>
</tr>
<tr>
<td>2.06</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Equ</td>
<td></td>
</tr>
<tr>
<td>2.13</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Equ</td>
<td></td>
</tr>
<tr>
<td>2.20</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Equ</td>
<td></td>
</tr>
<tr>
<td>2.26</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Pos</td>
<td>No change: 3 results are negative by IHC and positive by FISH (both cutoffs)</td>
</tr>
<tr>
<td>2.27</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Pos</td>
<td></td>
</tr>
<tr>
<td>2.32</td>
<td>Neg</td>
<td>Neg</td>
<td>Pos</td>
<td>Pos</td>
<td></td>
</tr>
<tr>
<td>8.44</td>
<td>Equ (2+) Neg</td>
<td>Pos</td>
<td>Pos</td>
<td>FISH-positive finding that might be overlooked by a 30% cutoff</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Equ, equivocal; IHC, immunohistochemistry; Neg, negative; Pos, positive.

* The IHC status is based on consensus of scores by 7 pathologists.

<table>
<thead>
<tr>
<th>HER2 Consensus Status, 30% Cutoff</th>
<th>Negative (0 and 1+)</th>
<th>Positive (3+)</th>
<th>Equivocal (2+)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (0 and 1+)</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>Positive (3+)</td>
<td>0</td>
<td>24</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Equivocal (2+)</td>
<td>5</td>
<td>0</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>95</td>
<td>24</td>
<td>18</td>
<td>137</td>
</tr>
</tbody>
</table>

**Table 7. Possible Change in Fluorescence In Situ Hybridization (FISH)–Positive Results in Human Epidermal Growth Factor Receptor 2 (HER2) Status by Changes in Cutoff Values**

**Table 8. Consensus of 7 Pathologists on Human Epidermal Growth Factor Receptor 2 (HER2) Status Using Manual Microscopy and the 10% and 30% Cutoffs**

(8% and the κ value was 0.91. An equivocal or weakly positive status by IHC was found in 17 samples using both cutoff values, and 3 of those samples were positive using FISH (FISH ratio of 2.33, 3.02, and 5.29, respectively). There were 3 IHC-negative-and-FISH-positive samples (FISH ratio of 2.26, 2.27, and 2.32, respectively) that both cutoff values classified as negative by IHC (Table 7). Additionally, the FISH status of 4 negative samples results by IHC changed from positive to equivocal when applying the ASCO/CAP FISH guidelines (Table 7).

**Interobserver Agreement Using Conventional Microscopy and Image Analysis-Assisted HercepTest Scoring**

Investigation of the interobserver agreement was done by pairwise comparisons of the 3 score groups (0 and 1+, 2+, and 3+) for all possible pathologist pairs. This was done for data obtained using conventional microscopy with the 10% cutoff and the 30% cutoff and with image analysis-assisted microscopy. For the 7 pathologists, 21 comparisons were possible. The mean (SD) interobserver κ value found in the scores using conventional microscopy and the Dako HercepTest guidelines was 0.72 (0.12) versus 0.84 (0.05) for image analysis-assisted microscopy and 0.72 (0.09) for conventional microscopy using the 30% cutoff. To test for a difference among these mean interobserver κ values, paired 2-tailed t tests were employed. The mean interobserver κ value for image analysis-assisted microscopy was significantly different from the means for conventional microscopy (P < .001), whereas the mean interobserver κ values found in conventional microscopy using 10% or 30% cutoff were not significantly different (P = .70). Therefore, interobserver agreement was not affected by the choice of cutoff value when scoring with a conventional microscope, whereas interobserver agreement was significantly better when pathologists used image analysis as an aid, compared with manual scoring.

**Concordance in HER2 Status Using FISH Scoring Is Similar With Conventional Microscopy and Image Analysis-Assisted HercepTest Scoring**

Agreement between HER2 status determined using image analysis-assisted scoring or conventional microscopy is shown in Table 4. When using image analysis-assisted microscopy with a cutoff of 10%, overall agreement was very similar (Table 4) to the overall agreement obtained by manual microscopy. A paired t test showed no significant difference between conventional microscopy and image analysis-assisted microscopy in concordance between HER2 status (positive/negative) obtained using HercepTest (10% cutoff) and HER2 FISH (2.0 cutoff; P = .70) when 2+ equivocal cases were excluded. This paired t test analysis was performed by comparing the 2 sets of κ values obtained from the stratified comparisons of conventional microscopy versus HER2 FISH and image analysis-assisted microscopy versus HER2 FISH (the averages of the κ values are shown in Table 4).

A difference in the scoring time for HercepTests was found among pathologists and between conventional microscopy and image analysis-assisted microscopy. The mean scoring times for the individual pathologists are plotted in the Figure. The mean (SD) scoring time (minute:seconds) was 1:01 (0:26) using conventional microscopy and 1:31 (0.38) using ACIS-assisted scoring. The range was from 0:17 to 2:22 for manual scoring and from 0:37 to 2:58 for image analysis-assisted microscopy. A 2-tailed paired t test of the mean scoring times revealed a significant difference (P < .001). For one pathologist, a shorter scoring time using image analysis was observed, and for another, no difference in time to score was observed.

**Comment**

The present study revealed a good concordance in the IHC scores among the 7 pathologists with both the 10% and a 30% cutoff values, and a consensus HER2 IHC assessment could be given to 140 (93%) and 138 (92%) of
The new guidelines have further been challenged\textsuperscript{24} by a group that considers FISH as the gold standard, suggesting FISH as the primary testing modality. One additional aspect of the ASCO/CAP guidelines that is not presently clear is the use of the “new” ASCO/CAP guidelines for the new FISH “equivocal” range. Because ASCO/CAP introduced an equivocal HER2 FISH range of 1.8 to 2.2, it is tempting to ask how a decision should be made in the clinic in cases of equivocal HER2 FISH (1.8–2.2). If performed by a highly experienced laboratory, repeated testing would be expected to lead to the same result. Importantly, our study revealed that the new ASCO/CAP guidelines might have resulted in the lack of appropriate treatment for at least one patient. That patient, with a FISH ratio of 2.0, had equivocal IHC results with the 30% cutoff and positive IHC results with the 10% cutoff. Therefore, the introduction of the new guidelines may leave this patient without a firm HER2 status because both methods changed the results from positive to equivocal. Even if the same sample is stained and scored several times with both IHC and FISH, a double-equivocal finding is possible (equivocal for both IHC and FISH). For patient care, the oncologist and the patient need a yes or a no answer, and they should be confident that the most appropriate therapy is being offered. Concern regarding whether a test is the most accurate or decisions about what to do with an equivocal result are problematic.

Our data also support the conclusion that there is only a limited difference in HER2 agreement between IHC and FISH when changing the cutoff levels because both the cutoff levels for IHC and FISH were increased. However, the use of the 30% cutoff raises the agreement to greater than the 95% concordance limit for positive and negative assay values\textsuperscript{25} because an increased number of 2+ equivocal cases are excluded. The cutoff of 30% was chosen to cover the term uniformly positive\textsuperscript{25} by the authors of the ASCO/CAP guidelines,\textsuperscript{3} but the cutoff is not related to a specific FISH ratio. The methods (IHC and FISH) are, however, not always expected to show identical results because different biological features are being measured (DNA or protein). Instead, the methods should be considered supplementary, meaning that a patient having either a HER2 3+ score or HER2 amplification should be considered as having a positive finding. Further, even if an equivocal zone for the FISH ratio can be an advantage as a guideline for retesting or additional counting, the rationale for moving the cutoff is lacking. If the cutoff is kept at 2.0 in accordance with clinical trial data there will be no difference in patient selection between the 10% and the 30% IHC cutoff, except for more FISH analyses being required because of the larger number of 2+ cases identified when the 30% cutoff is used for 3+ scores.

The time it takes to score a HercepTest result using image analysis-assisted microscopy is significant longer than it takes for manual scoring for most of the pathologists, although there were exceptions. However, the mean difference was about 30 seconds per slide. Additional observations made during this study lead to the conclusion that there is no observed difference in the concordance of HER2 status between HercepTest results and HER2 FISH pharmDx results when using conventional microscopy or image analysis-assisted microscopy. The observations showed that the use of image analysis did not impose an overall improvement or hindrance in the concordance of HER2 findings. In other words, the use of image analysis-assisted microscopy in this study did
not give better overall concordance with HER2 FISH status. This might simply reflect exceptional conventional HER2 scoring by the pathologists who scored the slides in this study (eg, there was a 92.1% overall agreement between FISH and manual scoring by the Dako guidelines, excluding 2+ cases, even though the pathologists had a range of experience). The benefit of digital microscopy to pathologist accuracy has been documented.10 Interobserver agreement was significantly higher using image analysis-assisted microscopy compared with conventional microscopy.

In conclusion, this study showed that a 30% cutoff value could mean higher concordance between the testing methods for 3+ cases. However, the 30% cutoff for 3+ IHC and 2.2 for HER2 FISH ratio could mean that fewer patients would be offered treatment with HER2-targeted therapy, especially if double equivocal HER2 results (both IHC and FISH) are obtained. The use of image analysis-assisted scoring increased the interobserver concordance.

We thank technologists Arman Vahebzadeh and Signe Lykke Nielsen for performing the HER2 FISH analyses and R. Lee Ryan, BS, for the data collection.

References