Evaluation of HER-2/neu Immunohistochemical Assay Sensitivity and Scoring on Formalin-Fixed and Paraffin-Processed Cell Lines and Breast Tumors

A Comparative Study Involving Results From Laboratories in 21 Countries

Anthony Rhodes, PhD,1 Bharat Jasani, PhD,2 Elizabeth Anderson, PhD,3 Andrew R. Dodson, FIBMS,4 and André J. Balaton, MD5

Key Words: HER-2/neu; Cell lines; Breast tumors; Immunohistochemistry; Assay sensitivity; Evaluation

Abstract

Variation in assay sensitivity was studied in more than 90 laboratories that assayed 4 formalin-fixed, paraffin-processed breast and ovarian carcinoma cell lines with graded levels of HER-2/neu protein overexpression and known levels of HER-2/neu gene amplification, in addition to breast carcinomas fixed and processed in the laboratories. Main methods were the HercepTest (DAKO, Ely, England) and individualized protocols using a polyclonal antibody and the CB11 clone. While the proportion of laboratories achieving appropriate results with the HercepTest was significantly higher than for participants using other assays, laboratories using other assays showed significant improvement in the second assessment run. The level of agreement in evaluations by 26 laboratories using the HercepTest was excellent on cell lines and tumors and was significantly greater than that achieved by the remaining 41 laboratories using other immunohistochemical methods. While laboratories using the DAKO HercepTest had the highest level of reproducibility in assay sensitivity and evaluation, the significant improvement in results by laboratories using other antibodies in the second assessment run suggests that stringent quality control and an ongoing quality assurance program using a standard reference material have the potential to improve the reliability of immunohistochemical assays for HER-2/neu, regardless of the antibody used.

Clinical trials have shown the potential benefits of trastuzumab (Herceptin) therapy for patients with invasive breast carcinomas that overexpress the HER-2/neu protein.1-3 Herceptin therapy, consisting of a humanized monoclonal antibody, is targeted at the HER-2/neu protein antigen and has the effect of inhibiting the growth of HER-2/neu-overexpressing tumor cells. To identify the 20% to 30% of women with breast cancer who will benefit most from trastuzumab therapy, a reliable and reproducible assay is required to detect HER-2/neu overexpression.4-6

Two main types of tests have evolved to identify such cases. Fluorescent in situ hybridization (FISH) is directed at detecting amplification of the HER-2/neu gene, which consistently is associated with HER-2/neu overexpression, while immunohistochemical analysis is used directly to detect overexpression of the HER-2/neu protein by the tumor cells.

The main advantage of the immunohistochemical assay over the FISH method is that it is faster and more economic to use, and the assay can easily be included as part of a routine diagnostic immunohistochemical service, currently offered by virtually all pathology laboratories. However, accurate assessment of immunohistochemical results is subject to variation of assay sensitivity and the method of evaluation used to interpret the results.7-12 Since Her-2/neu overexpression is considered predictive of a specific line of treatment, it is vital that the overall evaluation is accurate and based on a specific, reproducible immunohistochemical assay with appropriate sensitivity.

One way to ensure the accuracy and reproducibility of the results for HER-2/neu is by using robust internal quality control and external quality assessment systems. In this respect, external quality assessment has a role in determining
the level of interlaboratory agreement when conducting the same assay on the same material. This helps ensure that the same quality of service is provided to patients with breast cancer, regardless of where the person is treated.

Rhodes et al\textsuperscript{13} recently developed a quality control system consisting of 4 cell lines with differing levels of expression of HER-2/neu and with known HER-2\textsuperscript{neu} gene status, with the aim of providing a biologic standard against which HER-2/neu immunohistochemical assay sensitivity could be accurately evaluated, regardless of the method used. This standard subsequently was used to assess the variation in immunohistochemical assay sensitivity between laboratories from 21 countries when using different antibodies and methods to test for HER-2/neu overexpression on 2 separate occasions. The present article reports on the results of this study. In particular, it examines the level of agreement between a large number of centers in assay sensitivity when using different antibodies and the level of agreement between 67 of these laboratories and the assessors when evaluating the results with the widely used scoring system initially devised for the Clinical Trials Assay (CTA).\textsuperscript{2,14-16}

\section*{Materials and Methods}

\subsection*{Standard Cell Lines Block}

The cell lines used for the preparation of the standard cell lines block comprised the human breast carcinoma cell lines MDA-MB-453, BT-20, and MCF-7 and the ovarian carcinoma cell line SKOV-3. In a previous study, FISH analysis on these cell lines showed the SKOV-3 and MDA-MB-453 cell lines to have HER-2\textsuperscript{neu} gene amplification, while the cell lines BT-20 and MCF-7 did not.\textsuperscript{13} Briefly, the ratio of HER-2\textsuperscript{neu} to chromosome 17 signals were 4.2 and 3.0 for the SKOV-3 and MDA-MB-453 cell lines and 0.9 and 1.0 for the BT-20 and MCF-7 cell lines, respectively. A previously established method\textsuperscript{17} was used to prepare several identical formalin-fixed, paraffin-processed composite cell line blocks. Sections were cut from these blocks and placed on microscope slides according to the orientation depicted in Figure 1.

![Figure 1](https://academic.oup.com/ajcp/article-abstract/118/3/408/1758479)

\textbf{Figure 1} Orientation of cell lines on the distributed microscope slide.

\subsection*{Instructions to Participating Laboratories}

Laboratories participating in the United Kingdom National External Quality Assessment Scheme for Immunocytochemistry (UK NEQAS-ICC) at 2 consecutive assessment runs for HER-2/neu were provided with a slide (plus 1 spare) containing a paraffin section from the composite cell line block and were requested to do the following: (1) demonstrate HER-2/neu protein using their usual method and reagents; (2) at the same time as staining the cell lines to stain their own in-house tumor control using exactly the same method; (3) if using the HercepTest (DAKO, Ely, Cambridgeshire, England), to stain at the same time a DAKO cell line control from the HercepTest kit and submit this for assessment, along with the other slides; (4) describe precisely the method used to evaluate immunohistochemical assays for HER-2/neu; (5) evaluate the staining of the cell lines and their in-house tumor(s) using this scoring system; and (6) record their evaluations on the questionnaire and return it along with the stained slides to the UK NEQAS-ICC administrative center.

\subsection*{Assessment of Slides}

The immunohistochemical results were evaluated by 5 UK NEQAS-ICC assessors (the 5 authors) scoring independently and using the method of evaluation initially devised for the CTA.\textsuperscript{2,14-16} Briefly, the scoring system was as follows: no staining or membrane staining in fewer than 10% of tumor cells, 0; faint, barely perceptible membrane staining in more than 10% of tumor cells, 1+; weak to moderate complete membrane staining observed in more than 10% of tumor cells, 2+; and strong, complete membrane staining in more than 10% of tumor cells, 3+.

\subsection*{Procedure Following Assessment Runs}

After the 2 assessment runs, all participating laboratories received a report detailing the level of sensitivity achieved on the cell lines and in-house tumors, as evaluated by the UK NEQAS assessors. A detailed document, describing the cell lines and stating the assay sensitivity level most appropriate for the cell lines and details of the antibodies and assay methods used also was circulated to all participants.

\section*{Statistics}

The UK NEQAS-ICC assessment score was taken as the median value of the scores of the 5 UK NEQAS-ICC assessors evaluating the slides. The degree of agreement between the evaluations of the participating laboratories and those of the UK NEQAS-ICC assessors was assessed using the Cohen kappa statistic and 95% confidence intervals.\textsuperscript{18,19} The chi-square test was used to compare the differences in proportions of laboratories achieving appropriate results.
Results

Assay Sensitivity

A total of 94 laboratories from 21 countries (predominantly European) returned the cell line slides immunohistochemically stained for HER-2/neu in the first assessment run. In the second assessment run, 93 laboratories returned slides. A total of 78 laboratories participated in both assessment runs. The immunohistochemical assay scores achieved on the cell lines by these laboratories in the 2 assessment runs are summarized in Table I. The scores that most closely and consistently correlated with the known level of gene amplification of the cell lines were as follows: SKOV-3, 3+; MDA-MB-453, 2+; BT-20, 0 or 1+; and MCF-7, 0 or 1+. An overall total of 33 (35%) of 94 laboratories achieved this level of sensitivity in the first assessment run, and 58 (62%) of 93 achieved it in the second assessment run (chi square = 30.613; P < .001).

Choice of Antibody

Table II lists the antibodies used by participants and the assay sensitivity achieved with these markers by laboratories participating in both assessment runs. The HercepTest kit is listed separately from the polyclonal antibody A0485 owing to its use of a standard protocol. Both the kit and this polyclonal antiserum (with participants’ own individual protocols) were used by a large proportion of participants—65 (69%) of 94 in the first assessment run and 65 (70%) of 93 in the second assessment run. The CB11 clone was the other most commonly used HER-2/neu antibody, used by 20 (21%) of 94 participants in the first run and 19 (20%) of 93 in the second run. Other clones each accounted for 4% or less of the total returns in both runs.

Of laboratories participating in both assessment runs and using the HercepTest, 15 (68%) of 22 achieved the most appropriate staining pattern in the first assessment run, and 19 (76%) of 25 achieved it in the second assessment run. The proportion of participating laboratories that used their own individual protocols and achieved this pattern of staining with the DAKO polyclonal antibody A0485 was significantly lower for both runs, with 10 (29%) of 34 (chi square = 23.266; P < .001, 2-tailed) and 17 (55%) of 31 (chi square = 7.611; P = .012, 2-tailed) achieving appropriate staining, respectively. Similarly, for laboratories using clone CB11 (Novocastra, Newcastle upon Tyne, England), the proportions achieving appropriate staining on the first run (2/14 [14%]; chi square = 18.563; P < .001, 2-tailed) and the second run (6/15 [40%]; chi square = 10.658; P = .002, 2-tailed) were significantly lower than those achieved by laboratories using the HercepTest. However, there also was significant improvement in the number of laboratories achieving appropriate staining with all assays other than the HercepTest at the second assessment run compared with the results achieved for the first run (Table 2).

Level of Agreement Between Participating Laboratories Scores and UK NEQAS Scores at the First Assessment Run

A large proportion of participants (67/94 [71%]) used the CTA scoring system identical to the one used by the UK
NEQAS assessors. The remaining 27 participants used similar scoring systems based on various permutations of staining intensity and the proportion of cells showing membrane staining. The subsequent analysis was based on the evaluations of the 67 participants that clearly stated they used the CTA scoring system in the first assessment run. A total of 115 invasive breast carcinomas from 58 laboratories were returned with the completed questionnaires for evaluation. The staining of 12 tumors from 5 laboratories was considered by the UK NEQAS assessors to be impossible to interpret owing to excessive cytoplasmic or background staining; 4 of these tumors had been classified as unequivocally positive (scores of 3+) and 5 as unequivocally negative (scores of 0 or 1+). The 5 laboratories concerned also achieved inadequate results on the cell lines.

When all scores for each of the 4 cell lines stained by all the 67 participants using the CTA scoring system were compared with the scores of the UK NEQAS assessors, there was 74% total agreement on the cell lines (kappa, 0.64; 95% confidence interval, 0.57-0.71) and 73% total agreement on the breast tumors (kappa, 0.62; 95% confidence interval, 0.50-0.74). Agreement between just the 26 laboratories using the HercepTest and the UK NEQAS assessors was higher, with
86% of the evaluations being the same for the cell lines Table 3 and 87% for the breast tumors Table 4. Agreement between the scores of the UK NEQAS assessors and the scores of the remaining 41 laboratories using markers and methods other than the HercepTest but still using the CTA scoring system was significantly lower, with just 66% identical on the cell lines Table 5 (chi square = 55.285; \( P < .001 \)) and 62% on the tumors Table 6 (chi square = 31.875; \( P < .001 \)).

**Discussion**

Breast cancer is the most common form of cancer in women in the United Kingdom, with some 38,000 new cases diagnosed each year.\(^2^0\) The most practical and economic way to identify the 20% to 30% of women with breast cancer most likely to benefit from trastuzumab therapy would be by immunohistochemical analysis, but only if it can be shown that the assay for HER-2/neu is accurate and reproducible.

Quality assurance has an important role in ensuring that the results of these tests are of the same standard, irrespective of where a patient is treated. This ensures that the patient gets the same high quality treatment based on accurate tests and, in addition, ensures that the data accumulated on large patient populations from different institutions in clinical trials are reliable. To assess the accuracy with which the HER-2/neu immunohistochemical assays are currently applied in laboratories across the United Kingdom and in 20...
other countries, a biologically validated standard consisting of cell lines with a known and constant level of HER-2/neu expression was circulated to volunteer participant laboratories using different assay protocols.

An appropriate result in the present study was defined as the permutation of scores for the 4 cell lines (Image 1) that most closely correlated with their known level of gene amplification. When initially testing these cell lines, it became apparent that the MDA-MB-453 line, although having HER-2/neu gene amplification, could not be immunostained reliably with greater than 2+ staining intensity, without overstaining the BT-20 and MCF-7 cell lines (nonamplified), ie, these cell lines typically overstain as 2+ and 1+, respectively, if 3+ staining occurs on the MDA-MB-453 cell line. On tumors, an assay sensitivity level that equates to 1+ staining on the MDA-MB-453 cell line might have better correlation with FISH results, as many 2+ tumors have been shown by FISH not to have HER-2/neu gene amplification. However, provided guidelines are followed that recommend that 2+ staining on tumors be further analyzed by FISH, staining this cell line control and tumors with similar amounts of HER-2/neu expression as 2+ ensures that these tumors are further analyzed by FISH to identify which ones have HER-2/neu gene amplification and which ones do not. This, therefore, ensures that patients with 2+ tumors with gene amplification and patients with 2+ tumors without gene amplification all receive the appropriate therapy.

The proportion of laboratories achieving appropriate results with the HercepTest was significantly higher than for any other assay in both assessment runs. It is interesting to note that the proportion of laboratories achieving appropriate results with the HercepTest increased to more than 80% in both assessment runs, if only laboratories achieving appropriate results on the DAKO procedure control cell lines were included in the analysis, illustrating the close correlation that exists between appropriate results on both sets of cell lines. The results also show a dramatic and significant increase in the proportion of laboratories achieving an appropriate result on the cell lines in the second assessment run. However, the antibodies used by laboratories participating in both assessment runs were very similar, indicating that that improvement is not due to laboratories switching to the use of the HercepTest but due to laboratories improving their existing assays. Indeed, the only significant improvement seen between one assessment run and the next was noted with laboratories using assays other than the HercepTest (Table 2).
Arguably of equal importance to the technical reliability of the HER-2/neu assays is the reliability of the methods used to score the HER-2/neu results. A large proportion of the participating laboratories in the present study used a method of evaluation identical to that used in the HercepTest and to that used in the original CTA, with the vast majority of remaining laboratories using a very similar approach based on the intensity and the proportion of cells exhibiting membrane staining. Some studies have reported poor agreement between evaluators when evaluating the results of immunohistochemical assays for HER-2/neu using this method, which could account for some of the false-positive results, compared with the level of gene amplification reported in the literature, and which occur most commonly with tumors categorized as 2+. To study the reproducibility of the scoring systems used to determine the results of HER-2/neu assays, the use of a validated cell line standard in addition to tumor material has a number of advantages. First, it is possible to verify the level of gene amplification with FISH techniques on the

Arguably of equal importance to the technical reliability of the HER-2/neu assays is the reliability of the methods used to score the HER-2/neu results. A large proportion of the participating laboratories in the present study used a method of evaluation identical to that used in the HercepTest and to that used in the original CTA, with the vast majority of remaining laboratories using a very similar approach based on the intensity and the proportion of cells exhibiting membrane staining. Some studies have reported poor agreement between evaluators when evaluating the results of immunohistochemical assays for HER-2/neu using this method, which could account for some of the false-positive results, compared with the level of gene amplification reported in the literature, and which occur most commonly with tumors categorized as 2+. To study the reproducibility of the scoring systems used to determine the results of HER-2/neu assays, the use of a validated cell line standard in addition to tumor material has a number of advantages. First, it is possible to verify the level of gene amplification with FISH techniques on the

Table 4
Level of Agreement on In-House Breast Tumors Stained for HER-2/neu by HercepTest Users

<table>
<thead>
<tr>
<th>Participant Score</th>
<th>0</th>
<th>1+</th>
<th>2+</th>
<th>3+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK NEQAS Assessor Score</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>1+</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>2+</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>9</td>
<td>6</td>
<td>17</td>
<td>45</td>
</tr>
</tbody>
</table>

UK NEQAS, United Kingdom National External Quality Assessment Scheme.

The kappa score was 0.81 (95% confidence interval, 0.67-0.95). Total agreement 87%. For a description of the scoring system, see the “Materials and Methods” section. For proprietary information, see Table 3.

Table 5
Level of Agreement by Laboratories Using Immunohistochemical Assays Other Than the HercepTest on the Cell Lines SKOV-3, MDA-MB-453, BT-20, and MCF-7

<table>
<thead>
<tr>
<th>Participant Score</th>
<th>0</th>
<th>1+</th>
<th>2+</th>
<th>3+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK NEQAS Assessor Score</td>
<td>52</td>
<td>24</td>
<td>5</td>
<td>0</td>
<td>81</td>
</tr>
<tr>
<td>1+</td>
<td>2</td>
<td>8</td>
<td>11</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>2+</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>8</td>
<td>29</td>
</tr>
<tr>
<td>3+</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>36</td>
<td>35</td>
<td>39</td>
<td>164</td>
</tr>
</tbody>
</table>

UK NEQAS, United Kingdom National External Quality Assessment Scheme.

The kappa score was 0.53 (95% confidence interval, 0.43-0.63). Total agreement was 66%. For a description of the scoring system, see the “Materials and Methods” section. For proprietary information, see Table 3.

Table 6
Level of Agreement on In-House Breast Tumors Stained for HER-2/neu by Laboratories Using Immunohistochemical Assays Other Than the HercepTest

<table>
<thead>
<tr>
<th>Participant Score</th>
<th>0</th>
<th>1+</th>
<th>2+</th>
<th>3+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK NEQAS Assessor Score</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>1+</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>2+</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>3+</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>27</td>
<td>58</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>9</td>
<td>11</td>
<td>27</td>
<td>58</td>
</tr>
</tbody>
</table>

UK NEQAS, United Kingdom National External Quality Assessment Scheme.

The kappa score was 0.48 (95% confidence interval, 0.30-0.65). Total agreement was 62%. For a description of the scoring system, see the “Materials and Methods” section. For proprietary information, see Table 3.
whole cells as opposed to tissue sections. This makes the subsequent counts of signals representing the amplified genes easier and, therefore, more reliable. Second, the presence of misleading fixation and processing artifacts, frequently present in tissue samples, has been eliminated. Confusion about which is the most representative area to score therefore is avoided, as all the cells are fixed equally and evenly distributed. Third, cell lines are the material on which the CTA scoring system was initially established and currently are used as procedure controls to ensure the reliability of HercepTest staining before evaluation of the results on the tumor test material.\textsuperscript{15,16} Consequently, for the HercepTest, it is important to ensure the reproducibility of the scoring system on cell lines and tissue samples, as the evaluation of both ultimately will influence the reliability of the results. If there is a weakness in the scoring system, comparison of the evaluations performed by laboratories on cell lines and tumor samples may help identify whether the problem lies in the heterogeneity of the tumors and their fixation and processing or purely in the difficulties encountered in accurately defining cell membrane staining intensity using a 4-point scoring system.

Development by DAKO of the HercepTest and its approval by the US Food and Drug Administration (FDA) introduced a new approach to clinical pathology, one that attempts to prospectively standardize methods used by insisting that users of the HercepTest adhere to strict guidelines on how the tissue is fixed, through the use of a detailed technical method and interpretation of the results. Individual laboratories are “trained” in the technical aspects of the FDA-approved method and in evaluating the results using the CTA scoring system. One would expect, therefore, that the stringent training program and guidelines imposed would result in a greater level of reproducibility between different sites, not only in assay sensitivity but also in the evaluation of the results compared with those from centers not subjected to this stringent training program. Part of the present study tested this hypothesis by comparing the results of participating laboratories using the HercepTest with those from other laboratories not using this kit.

The results showed that the agreement between the UK NEQAS assessors and participants using the HercepTest was excellent, both on the cell lines and the in-house tumors. The only disagreements in the evaluation of in-house cases was the recording of 3 tumors as 1+ that were considered to have 0 staining by the UK NEQAS assessors and the scoring of 3 tumors as 2+ that were scored as 1+ by the assessors (Table 4). Assuming guidelines are implemented, recommending that all cases recorded as 2+ by immunohistochemical assays represent an equivocal result that requires confirmation of HER-2/neu gene amplification by molecular methods such as FISH,\textsuperscript{14} then none of these disagreements is likely to have resulted in the misclassification of patients. On the cell lines, the only serious discrepancy in the evaluations was that 3 cell lines considered to be 2+ by the UK NEQAS assessors were scored as 1+ (n = 1) and 3+ (n = 2) by the participants. This accounts for just 3% of the evaluations on the cell lines by laboratories using the HercepTest (Table 3). In comparison, the agreement between the evaluation of scores by the UK NEQAS assessors and the participants using immunohistochemical assays other than the HercepTest was poor, and the proportion of evaluations showing complete agreement was significantly lower on both tumors and cell lines (Tables 5 and 6). This discordance would have resulted in tumors being erroneously reported as being unequivocally positive or unequivocally negative in 10 (17%) of 58 cases and serious discrepancies in 12 (7.3%) of the cell lines (n = 164 samples).

The level of agreement between the participants and the UK NEQAS assessors in evaluating the cell lines tended to reflect the level of agreement recorded for the evaluation of the in-house tumors. Consequently, the results of this study suggest that lack of reproducibility of the CTA scoring system between laboratories is not primarily due to the heterogeneity of the tumors being evaluated or the different ways in which they have been fixed and processed but due to how stringently the scoring criteria are applied. It has been shown that the level of agreement between the UK NEQAS assessors and laboratories using the HercepTest and, therefore, having undergone stringent training in the application of the CTA scoring system was significantly greater than the level of agreement between the UK NEQAS assessors and laboratories using a variety of other immunohistochemical methods. This suggests that when the CTA scoring system is adhered to strictly, it gives excellent agreement in scoring between different laboratories, both in the evaluation of tumors and in the evaluation of cell lines stained for HER-2/neu. Future work therefore should focus on promoting the use of a standard scoring system and introducing educational programs to ensure that the criteria defining the scoring system are strictly followed.

Once educational programs are able to ensure that the scoring system is applied in a highly reproducible manner, the technical issues relating to interlaboratory variation in assay sensitivity and reproducibility will need to be addressed. The great value of the results reported in the present study on the reproducibility of the HercepTest is that they show that standardization of immunohistochemical analysis is an achievable concept. However, the results also show that laboratories using other antibodies to HER-2/neu with their own customized assays show significantly improved results at a second run of a quality assurance program using a cell line standard. This suggests that stringent quality control and an ongoing quality assurance
program using a standard reference material and an intensive educational program have the potential to provide a level of standardization equivalent to or greater than that currently achieved by laboratories using the HercepTest, regardless of the antibody used. Standardization by this approach may well permit the FDA to broaden its certification of additional HER-2 reagents, permitting laboratories to choose from various commercial suppliers and manufacturers of specific reagents. This, in turn, will keep costs down and not exclude antibodies and desirable technical innovations that may further improve the reliability of immunohistochemical tests. In this respect, it is important that assays other than the HercepTest continue to be validated, as markers such as the CB11 and TAB 250 clones have recently been shown to have a greater level of concordance with HER-2/neu gene amplification as measured by FISH and to have greater statistical significance with respect to patient response rates to combined trastuzumab and paclitaxel therapy.

From the 1 UK NEQAS-ICC and the Department of Histopathology, University College London Medical School, London, England; 2 Immunocytochemistry and Molecular Pathology Unit, University of Wales College of Medicine, Cardiff; 3 Clinical Research Department, Christie Hospital NHS Trust, Manchester, England; 4 Molecular Pathology and Development, Department of Pathology, Royal Liverpool University Hospital, Liverpool, England; and 5 Centre de Pathologie, Bievres, France.

Address reprint requests to Dr Rhodes: Dept of Histopathology, University College London Medical School, Rockefeller Bldg, University St, London WC1E 6JJ, England.

Acknowledgments: We thank all participating laboratories without whose support this study would have been impossible and Simon Cross, MD, Department of Pathology, University of Sheffield Medical School, Sheffield, England, for advice on the use of kappa statistics.

References


