

## Best Practice No 176

# Updated recommendations for HER2 testing in the UK

I O Ellis, J Bartlett, M Dowsett, S Humphreys, B Jasani, K Miller, S E Pinder, A Rhodes, R Walker

This paper serves to update previously published guidance on rationale and methodology for HER2 laboratory testing following the recommendation for the use of HER2 targeted treatment in the management of advanced breast cancer in the UK. Emphasis is placed on the standardisation of methodology and assessment and strategies to achieve high quality performance. A two phase testing algorithm based on first line immunocytochemistry evaluation and second line fluorescence in situ hybridisation assessment of borderline cases is recommended. To ensure maintenance of expertise, an annual caseload volume of at least 250 cases is recommended for laboratories providing a testing service.

> he humanised anti-HER-2/neu (also known as c-erbB-2, further denoted HER2) monoclonal antibody trastuzumab (Herceptin),<sup>1</sup> has recently been endorsed by the UK National Institute for Clinical Excellence (NICE) for the treatment of metastatic disease,2 (fig 1) and establishing tumour HER2 status is a prerequisite for the use of trastuzumab.<sup>1</sup><sup>2</sup> These guidelines update the previous UK HER2 testing guidelines.<sup>3</sup> and have been formulated to give advice on methodology and quality assurance for local testing to ensure that HER2 testing results are accurate and reliable, regardless of the test that is used. Authorship is based on representation of existing UK groups with experience in HER2 assay methodology and performance, quality assurance, guideline production, and establishing performance standards. These guidelines aim to offer guidance to those laboratories that are considering offering an HER2 evaluation service, provide information on internal and external quality assurance, and provide a basis for establishing service standards for use by professional and National Health Service groups in the UK. NICE has referred to the need to adhere to our published guidelines and this document serves to update the existing guidance.

See end of article for authors' affiliations

Correspondence to: Dr I O Ellis, Department of Histopathology, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK; ian.ellis@nottingham.ac.uk

Accepted for publication 3 June 2003

"These guidelines update the previous UK HER2 testing guidelines, and have been formulated to give advice on methodology and quality assurance for local testing to

J Clin Pathol 2004;57:233-237. doi: 10.1136/jcp.2003.007724

ensure that HER2 testing results are accurate and reliable, regardless of the test that is used"

## GENERAL PRINCIPLES Suitable samples

Formalin fixed, paraffin wax embedded tumour tissue samples are appropriate for assay.<sup>4-9</sup> Ideally, buffered formalin should be used for fixation because the use of Bouin's fixative will preclude testing by fluorescence in situ based methods. Other methods of tissue fixation can also adversely affect antigen reactivity.

## Caseload

- Laboratories providing a testing service should be carrying out a minimum of 250 assays each year for immunohistochemical detection of HER2. This target value has been set to ensure continuing expertise of assay providers. There is evidence of higher consistency of assay quality when tests are performed by high volume reference laboratories.<sup>10</sup> <sup>11</sup> Specifically, laboratories with low volumes using immunohistochemistry (IHC) have higher frequencies of IHC positive cases and reduced levels of IHC/FISH (fluorescence in situ hybridisation) concordance, which is related to a higher proportion of IHC positive/FISH negative results.
- Centres with low numbers of cases (< 250/ year) requiring IHC assay should consider using a reference laboratory service.
- Similar principles apply to FISH testing and it is recommended that laboratories testing fewer than 100 cases each year should consider referring their workload to a reference laboratory. A smaller caseload has been set for the FISH assay because it is generally accepted to be a more discriminant test at the positive negative borderline, has greater ease of methodological standardisation, and less observer variation

## Appropriate laboratory assay methods

IHC and FISH<sup>5-9</sup> are the techniques recommended for determining HER2 status. Currently, the

Abbreviations: FISH, fluorescence in situ hybridisation; IHC, immunohistochemistry; NICE, National Institute for Clinical Excellence; UK NEQAS-ICC, United Kingdom National External Quality Assessment Scheme for Immunocytochemistry

#### NICE summary statement

The clinical effectiveness and cost effectiveness of trastuzumab for the treatment of advanced breast cancer

#### 1 Guidance

- 1.1 Trastuzumab in combination with paclitaxel (combination trastuzumab is currently only licensed for use with paclitaxel) is recommended as an option for people with tumours expressing human epidermal growth factor receptor 2 (HER2) scored at levels of 3+ who have not received chemotherapy for metastatic breast cancer and in whom anthracycline treatment is inappropriate
- 1.2 Trastuzumab monotherapy is recommended as an option for people with tumours expressing HER2 scored at levels of 3+ who have received at least two chemotherapy regimens for metastatic breast cancer. Prior chemotherapy must have included at least an anthracycline and a taxane where these treatments are appropriate. It should also have included hormonal therapy in suitable cestrogen receptor positive patients
- 1.3 HER2 levels should be scored using validated immunohistochemical techniques and in accordance with published guidelines. Laboratories offering tissue sample immunocytochemical or other predictive tests for treatment response should use validated standardised assay methods and participate in and demonstrate satisfactory performance in a recognised external quality assurance scheme

Figure 1 Updated National Institute for Clinical Excellence (NICE) recommendations for the treatment of metastatic disease by trastuzumab (Herceptin).

other available HER2 testing techniques (chromogenic in situ hybridisation, polymerase chain reaction, enzyme linked immunosorbent assay, and Southern blotting) should be used for research only.

For both immunohistochemical and FISH based HER2 testing, comprehensive standardisation of methodology, including monitoring of scoring procedures, and the inclusion of validated controls, is mandatory. In the UK, participation and satisfactory performance in the current NEQAS scheme for IHC and the forthcoming NEQAS scheme for HER2 FISH is a requirement. These schemes are open to laboratories across Europe. Although published data support the use of FISH for the selection of patients most likely to respond to trastuzumab, the current UK licence for this agent allows the treatment of patients with tumours strongly staining by IHC. Worldwide, there remains an ongoing debate as to whether laboratories should switch to the use of FISH for all specimens, removing the need for a second tier of testing to identify HER2 positive cases, or adopt the two tier testing strategy (fig 2) currently in use in the UK

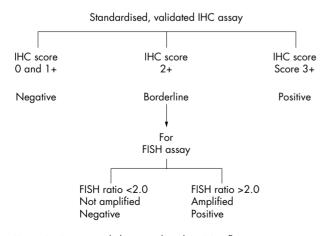


Figure 2 Recommended testing algorithm. FISH, fluorescence in situ hybridisation; IHC, immunohistochemistry.

reference laboratories. Current experience from the UK reference laboratories indicates that there is a very high level of correlation between IHC and FISH assay results in the 0/1+ and 3+ IHC categories, negating the need for dual IHC and FISH based assays in most cases<sup>12</sup>; however, other published studies show higher rates of discordance in these categories. Caution may be needed before extrapolating the experience of the reference centres to laboratories with lower case loads, and laboratories are strongly advised to consider an initial validation testing phase to confirm high levels of concordance between IHC and FISH for IHC 0/1+ cases and 3+ cases.

Although the UK licence remains focused on IHC positivity it is logical, in the light of such data, to use FISH as a secondary test in the equivocal 2+ IHC category to clarify the HER2 status of these cases (fig 2); however, once trastuzumab is licensed for both FISH and IHC positive cases it is possible that any advantage of the current two tier testing system will be scrutinised. In this case, as at present in other counties, some laboratories will choose to use FISH as a front line diagnostic test without the use of IHC. It is also expected that emerging data on accuracy of prediction of response to HER2 targeted treatments will influence the choice of testing method.

In summary, current UK recommendations are for a two tier testing strategy, using the model show in fig 2, but this does not preclude laboratories from using primary FISH testing once there is drug licence revision of trastuzumab to include FISH positive status as an indication for treatment.

#### Controls

The inclusion of controls and their detailed scrutiny are essential to ensure test accuracy. A recommended positive control or controls producing results close to important decision making points and a negative control are recommended.

Cell line preparations containing multiple samples of known HER2 status characterised by FISH and IHC are useful as controls.<sup>13</sup> Where possible, tissue based controls, preferably from breast cancers, should also be used in all assay runs.

Excessive antigen retrieval can be monitored by an evaluation of normal breast epithelial cells as an internal control. Should membrane reactivity be identified in the normal cell population, excessive antigen retrieval may have occurred and retesting of the entire run should be considered. It is preferable to titrate the staining methodology to avoid detecting normal degrees of HER2 expression, rather than advocate a scoring system based on subtracting the degree of staining seen in normal cells from tumour cell staining.

#### **Evaluation**

For the assessment of both IHC and FISH preparations training and experience in the interpretation of histological characteristics of breast tissue are essential. The recognition of different histological tumour types is required. In particular, the HER2 status should be determined only on the invasive portion of the tumour and IHC positivity or FISH amplification should not be reported as a positive result in isolation. Image analysis systems are currently under investigation and may provide alternatives to manual scoring for both IHC and FISH in the future; however, at present, insufficient evidence is available to recommend their routine use in the diagnostic setting.

#### IMMUNOHISTOCHEMISTRY For all IHC tests

Antigen retrieval processes are crucial; they must be standardised and follow strict protocols. The antibody used and its titre should be predefined. Standardisation can be achieved using commercial assay systems such as the Herceptest (DakoCytomation, Ely, Cambridgeshire, UK). For in house assays no single antibody has consistently been shown to be superior in terms of specificity and sensitivity. At present, antibody clones CB11 (Novocastra, Newcastle upon Tyne, UK), TAB 250 (Zymed, San Francisco, California, USA), and polyclonal anti-sera AO485 (DakoCytomation) are the most widely used for all assay methods. Test conditions (temperature, exposure time, etc) should be standardised. To achieve good concordance between IHC and FISH assays it is recommended that the sensitivity of the IHC method should be set below that which detects normal amounts of HER2 protein in benign or normal breast epithelial cells.

#### Validation of standardised assay method

Test conditions should be optimised so that distinct moderate or strong membrane staining identifies FISH positive samples. This can be achieved by:

(1) Dual IHC and FISH assay of a contemporary series of breast carcinomas (minimum 100 cases). FISH can be confined to those cases demonstrating some evidence of membrane reactivity (1, 2, or 3+).

(2) Laboratories can use a series of recent cases from their own hospital. Alternatively, the use of tumour tissue array blocks for validation may reduce costs. It may be possible to obtain such sections, which have already been scored for IHC and FISH, from a research laboratory or reference source.

(3) Laboratories should aim to achieve high degrees of concordance between IHC 0/1+ results and FISH negativity and between IHC 3+ results and FISH positivity.

Laboratories not wishing to standardise in house methodology should consider using a commercial kit assay system such as the Herceptest (DakoCytomation).

#### Scoring IHC

Only membrane staining of the invasive tumour should be considered when scoring IHC tests. If a commercial kit assay system is used, it is recommended that laboratories adhere strictly to the kit assay protocol and scoring methodology. Local modifications of techniques can lead to false positive and negative assay results. The scoring method recommended is a semiquantitative system based on the intensity

Score to report	HER2 protein overexpression assessment	Staining pattern
0	Negative	No staining is seen, or membrane staining in less than 10% of tumour cells
1+	Negative	A faint/barely perceptible membrane staining is detected in more than 10% of tumour cells. The cells are only stained in part of the membrane
2+	Borderline	Weak to moderate complete membrane staining is seen in more than 10% of tumour cells
3+	Positive	Strong complete membrane staining is seen in more than 10% of tumour cells

Figure 3 Recommended immunohistochemistry scoring method.

of the reaction product and the percentage of membrane positive cells, giving a score range of 0 to 3+ (fig 3). Samples scoring 3+ are regarded as unequivocally positive and 0/1+ as negative. Borderline 2+ samples require confirmation using another analysis system, ideally FISH (fig 2). Some commercial providers of kit assay systems provide detailed guidance on scoring IHC assays, which can help observers to avoid pitfalls and problems. Observers should be aware of the range of common artefacts, including edge artefacts, which can be problematic in small biopsy samples, and the effects of variation in method sensitivity, such as excessive antigen retrieval, leading to background staining and normal cell membrane reactivity.

#### "Only membrane staining of the invasive tumour should be considered when scoring immunohistochemistry tests"

Non-commercial kit assay methods can be scored on a similar basis or by modification to a three tier system of positive, borderline, and negative. Until better evidence on scoring methodology emerges, the cutoff points for such simplified assay scoring systems should be based on the existing HercepTest kit method, with a positive score being based on a 3+ score, borderline on a 2+ score, and negative a 1+ or 0 score (figs 2 and 3).

Interobserver variation in the assessment of staining can lead to misclassification of HER2 status.<sup>14</sup> Each individual assessor should standardise scoring against known positive, negative, and borderline cases. It is also preferable to assess comparability of scoring with a colleague on a regular basis.

#### Quality assurance

All clinical laboratories using assays for HER2 as predictive or prognostic tests must participate in an appropriate external quality assurance (EQA) programme, such as that run by the United Kingdom National External Quality Assessment Scheme for Immunocytochemistry (UK NEQAS-ICC).

On a quarterly basis, UK NEQAS-ICC circulates unstained sections from a formalin fixed and paraffin wax processed block comprising human breast carcinoma cell lines (MDA-MB-453, BT-20, and MCF-7) and an ovarian carcinoma cell line (SKOV-3) to over 100 laboratories. Previous FISH analysis on these cell lines has shown the SKOV-3 and MDA-MB-453 cell lines to have HER2 gene amplification, whereas the BT-20 and MCF-7 cell lines do not.13 When the sensitivity of the assay is appropriate, the cell line SKOV-3 should be unequivocally positive (3+) and the cell lines MCF-7 and BT-20 should be unequivocally negative (0 or 1+). The most appropriate result on the cell line MDA-MB-453 is 2+. Following strict adherence to the Dako HercepTest staining protocol, it has been shown that over 80% of laboratories using the HercepTest achieve this permutation of immunostaining on the cell lines SKOV-3, MDA-MB-453, BT-20, and MCF-7. Laboratories using individually customised assays utilising the CB11 and TAB 250 clones and DAKO polyclonal antisera have achieved equivalent staining.

Participating laboratories are requested to test the UK NEQAS sections and their own in house control for HER2 and to return them to the organising centre for evaluation by a panel of five expert assessors using the method of evaluation initially devised for the Herceptin clinical trials assay, with the median value from the five assessors being taken as the final score.<sup>13–15</sup>

To identify and rectify suboptimal performance for HER2 assays by UK laboratories within an acceptable time frame, UK NEQAS-ICC will approach all UK laboratories achieving an inappropriate result on the UK NEQAS sections (a score other than 3+, 2+, 0/1+, 0/1+ on the cell lines SKOV-3, MDA-MB-453, BT-20, and MCF-7, respectively) and provide

advice for improvement. Any of these participants subsequently achieving an inappropriate result at two subsequent assessment runs on the UK NEQAS sections will be issued a warning letter. With the issue of this warning letter, UK NEOAS will provide further technical advice and support. including attendance at the UK NEQAS organisers' laboratory by biomedical scientists from the poor performing laboratory. All attempts will be made to assist the laboratory to improve. However, failure to do so, with the laboratory accruing a total of four successive inappropriate scores on the UK NEQAS sections, despite intensive advice and assistance, will result in the laboratory concerned being removed from the UK NEQAS for HER2 scheme register and being reported to the chairman of the National Quality Assurance Advisory Panel (NQAAP). This may ultimately result in the laboratory concerned losing its CPA status for this test. However, the laboratory will be permitted to continue participating in EQA for HER2 (if it so wishes) and the chairman of NQAAP will be notified if it is able to show significant improvement by subsequently accruing acceptable results at all of four successive assessment runs. This approach will ensure that poorly performing laboratories are identified promptly and the situation rectified through appropriate action being taken within a 12 month period, either by the laboratories showing improvement to an acceptable standard or by being removed from the UK NEQAS participation register and loosing their accreditation status for this test.

#### FLUORESCENCE IN SITU HYBRIDISATION (FISH)

FISH testing for HER2 should meet the following criteria:

(1) Comprehensive standardisation of methodology.

(2) Validated controls. The inclusion of a chromosome 17 control to allow for correction of the HER2 signal number for chromosome 17 aneusomy (seen in over 50% of cases) is considered beneficial by many laboratories and is recommended.

(3) Validated scoring procedures.

#### **General principles**

There is no evidence that storage of blocks or slides leads to the deterioration of signal. However, it is recommended that storage of cut sections from controls or samples for over six to 12 months should be avoided.

It is advisable to locate areas of invasive tumour using a serial section stained with haematoxylin and eosin and to use this to locate tumour areas to be scored after testing. Care should be taken to avoid areas of ductal carcinoma in situ, which can show amplification even when adjacent invasive tumour cells are negative. With experience, such features can be identified under fluorescence microscopy; however, the use of serial haematoxylin and eosin sections is essential should there be any uncertainty in this area.

Tissue digestion should be standardised to maintain nuclear morphology and should follow strict protocols.<sup>16</sup> Some laboratories find it helpful to evaluate nuclear structure before hybridisation and adjust digestion, where appropriate, to preserve nuclear integrity. This may be particularly valuable with difficult sections, cytology samples, bone biopsies, etc. Evaluation of sections before hybridisation can also improve efficiency and is recommended. Hybridisation and washing steps should be standardised. Guidance can be provided by the reference laboratories. Use of automated tissue processors and standardised commercial tissue digestion kits can improve consistency and should be considered.

It is recommended that commercially available probes are used. For systems using in house nick translated probes,

attention should be given to batch variability of nick translation enzymes, and other reagents.

Laboratories not wishing to use in house methods should consider using a commercial system such as Vysis PathVysion (Abbott UK, Queensborough, Kent, UK). Other commercial systems currently available are not yet widely validated or lack the chromosome 17 control discussed above.

#### Scoring FISH

HER2 FISH testing results are conventionally expressed as the ratio of HER2 signal to chromosome 17 signal. Tumours showing a ratio of more than 2 should be considered positive. Cut off values for HER2 gene amplification when chromosome 17 probes are not used have not been established.

## "Published data suggest that interobserver variation is significantly lower for fluorescence in situ hybridisation than for immunohistochemistry"

The number of chromosome 17 and HER2 signals is scored for between 20 and 60 cells where possible, using at least three distinct tumour fields, and the mean HER2 to chromosome 17 copy ratio is calculated. In most cases, where either clear amplification is observed or the ratio is below 1.5, scoring of 20 cells is sufficient. In cases where either tumour heterogeneity is seen (1-2% of cases) or the ratio is close to 2.0 (between 1.5 and 2.3) more cells should be scored (up to 60). Samples with more than 2.0 copies of HER2 for each chromosome 17 are considered to be amplified. Published data suggest that interobserver variation is significantly lower for FISH than for IHC.<sup>17</sup> In this study, interobserver variations for both FISH and IHC analyses were evaluated on the same patient cohort, and, as measured by  $\kappa$ statistics, were lower for two IHC based tests (0.67 and 0.74) than for FISH testing (0.973; 17). Using IHC based testing, concordance between two observers in selecting patients for 0/1 + v 2/3 + intensity was seen in 81% of cases, whereas for FISH, concordance in the identification of amplified and nonamplified cases was greater than 99%.17 Further data on reproducibility of IHC and FISH scoring are reviewed elsewhere.<sup>18 19</sup> However, especially when developing a new service, care needs to be taken. The recommendation is that laboratories should perform validation studies by dual observer scoring when training new staff until interobserver variation for normal specimens and those with low level amplification is routinely below 15%. Continued monitoring of scoring offers advantages in quality control and training but is not a requirement. Variation increases with highly amplified samples, and is not crucial where the ratio exceeds 4

#### Quality assurance

To ensure adequate quality assurance, UK laboratories wishing to set up independent FISH testing are recommended to join the proposed EQA scheme coordinated by NEQAS. Currently we envisage using tissue microarrays or multiblocks to provide adequate material for analysis. The scheme will be designed to evaluate methodological and scoring aspects but may not cover morphological aspects.

## Authors' affiliations

I O Ellis, S E Pinder, Department of Histopathology, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK (National Coordinating Group for Breast Pathology and National HER2 Reference Laboratory)

J Bartlett, University Department of Surgery, Glasgow Royal Infirmary, Glasgow G4 OSF, UK (National HER2 Reference Laboratory Service) M Dowsett, Department of Biochemistry, Royal Marsden Hospital, London SW3 6JJ, UK (National HER2 Reference Laboratory Service) S Humphreys, Department of Cellular Pathology, Preston Hall Hospital, Maidstone, Kent ME20 7NJ, UK (National Coordinating Group for Breast Pathology)

B Jasani, Department of Pathology, University of Wales College of Medicine, Cardiff CF14 4XN, UK (United Kingdom National External Quality Assessment Scheme for Immunocytochemistry) K Miller, A Rhodes, Department of Histopathology, University College London Medical School, London WCIE 6JJ, UK (United Kingdom

National External Quality Assessment Scheme for Immunocytochemistry) R Walker, Department of Pathology, University of Leicester, Leicester, UK (National Coordinating Group for Breast Pathology)

#### REFERENCES

- Plunkett TA, Miles DW. New biological therapies for breast cancer. Int J Clin Pract 2002;56:261-6
- National Institute for Clinical Excellence. Technology Appraisal No. 34. 2 Guidance on the use of trastuzumab for the treatment of advanced breast cancer 2002 (http://www.nice.org.uk/article.asp?a=29280)
- 3 Ellis IO, Dowsett M, Bartlett J, et al. Recommendations for HER2 testing in the UK. J Clin Pathol 2000;53:890-2.
- 4 Press MF HG, Godolphin W, et al. Sensitivity of HER2/neu antibodies in archival tissue samples: potential source of errors in immunohistochemical studies of oncogene expression. *Cancer Res* 1994;**54**:2771–7. **Walker R**. The significance of histological determination of HER-2 status in
- 5 breast cancer. Breast 2000;9:130-3.
- 6 Schnitt SJ. Breast cancer in the 21st century: Neu opportunities and neu challenges. Mod Pathol 2001;14:213-18.
- Wang S, Saboorian MH, Frenkel E, et al. Laboratory assessment of the status of Her-2/neu protein and oncogene in breast cancer specimens: comparison of immunohistochemistry assay with fluorescence in situ hybridisation assays. J Clin Pathol 2000;53:374-81

- 8 Lehr H-A, Jacobs TW, Yaziji H, et al. Quantitative evaluation of HER-2/neu status in breast cancer by fluorescence in situ hybridization and by immunohistochemistry with image analysis. Am J Clin Pathol 2001;115:814-22.
- 9 Rampaul RS, Pinder SE, Gullick WJ, et al. HER-2 in breast cancer-methods Oncol Heratol 2002;**43**:231–44.
- Oncon Function 2002 (2012) 10 Paik S, Bryant J, Tan-Chiu E, et al. Real-world performance of HER2 testing national surgical adjuvant breast and bowel project experience. J Natl Cancer Inst 2002;94:852-4.
- 11 Roche PC, Suman VJ, Jenkins RB, et al. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. J Natl Cancer Inst 2002;94:855–7.
- 12 Dowsett M, Ellis IO, Bartlett JMS, et al. Correlation between immunohistochemistry and FISH for HER-2 in 441 breast carcinomas from multiple hospitals. J Pathol 2001;195(suppl):5A.
- Rhodes A, Jasani B, Couturier J, et al. A formalin-fixed, paraffin-processed cell line standard for quality control of immunohistochemical assay of HER-2/neu expression in breast cancer. Am J Clin Pathol 2002;117:81-9.
- 14 Rhodes A, Jasani B, Anderson E, et al. Evaluation of HER-2/neu immunohistochemical assay sensitivity and scoring on formalin fixed and paraffin processed cell lines and breast carcinomas: a comparative study involving results from laboratories in 21 countries. Am J Clin Pathol 2002;**118**:408–17.
- 15 Mass R. The role of Her-2 expression in predicting response to therapy in breast cancer. Semin Oncol 2000;27:46-52
- 16 Watters AD, Bartlett JM. Fluorescence in situ hybridization in paraffin tissue
- sections: pretreatment protocol. Mol Biotechnol 2002;21:217-20.
  Bartlett JMS, Going JJ, Mallon EA, *et al.* Evaluating HER2 amplification and overexpression in breast cancer. J Pathol 2001;195:422-8.
  Bartlett JMS, Mallon EA, Cooke TG. The clinical evaluation of HER2 status,
- which test to use? J Pathol 2003;199:411-17.
- 19 Bartlett JMS, Mallon EA, Cooke TG. Molecular diagnostics for determination of HER2 status. Diagn Pathol 2003;9:48-55.