

Standardization of HER2 testing: results of an international proficiency-testing ring study

Mitch Dowsett¹, Wedad M Hanna², Mark Kockx³, Frederique Penault-Llorca⁴, Josef Rüschoff⁵, Thorsten Gutjahr⁶, Kai Habben⁷ and Marc J van de Vijver⁸

¹Academic Department of Biochemistry, Royal Marsden Hospital, London, UK; ²Department of Anatomic Pathology, University of Toronto, Toronto, ON, Canada; ³Histogenex, Antwerp, Belgium; ⁴Département de Pathologie, Centre Jean-Perrin, Clermont-Ferrand, France; ⁵Institute für Pathologie, Klinikum Kassel, Kassel, Germany; ⁶F Hoffmann-La Roche Ltd, Basel, Switzerland; ⁷Roche Diagnostics GmbH, Penzberg, Germany and ⁸Department of Pathology, Netherlands Cancer Institute, Amsterdam, The Netherlands

Human epidermal growth factor receptor 2 (HER2) positivity in breast cancer is a prognostic factor regarding tumor aggressiveness and a predictive factor for response to trastuzumab (Herceptin®). Early and accurate HER2 testing of all breast cancer patients at primary diagnosis is essential for optimal disease management. Routine HER2 tests, such as immunohistochemistry and fluorescence *in situ* hybridization (FISH), are subject to interlaboratory variation, and validation by laboratory proficiency testing is important to improve standardization. This study compared immunohistochemistry and FISH testing between five international pathology reference centers. Each center evaluated 20 immunohistochemistry and 20 FISH breast cancer specimens in five testing rounds. In each round, one center selected two sets of four different invasive tumor specimens (set A for immunohistochemistry and set B for FISH) and sent samples to the other four centers in a blinded manner, while retaining samples for its own evaluation. Results were analyzed by an independent coordinator. With immunohistochemistry, there were no differences between the five centers for any of the specimens at the level of diagnostic decision (positive or negative HER2 status). However, differences between laboratories were observed in immunohistochemistry scoring. Of the 20 specimens, four were scored as negative (0/1+) and five as positive (3+) in all centers; eight were negative or equivocal (2+), and three positive or equivocal. After FISH retesting of nine of the 11 equivocal immunohistochemistry cases, consensus was achieved in 15 of 18 (83%) specimens. FISH analysis of set B specimens resulted in consensus between centers in 16 of 20 (80%) specimens (six negative and 10 positive). All four discordant FISH specimens were scored as having *HER2:CEP17* ratios within the range 1.7–2.3 by at least one center. Equivocal immunohistochemistry and borderline FISH cases are difficult to interpret, even for highly experienced and validated laboratories, highlighting the need for quality-control procedures.

Modern Pathology (2007) 20, 584–591. doi:10.1038/modpathol.3800774; published online 30 March 2007

Keywords: breast cancer; HER2; quality control; trastuzumab

Recent estimates place the lifetime risk of developing breast cancer at 8 and 13% for women living in Europe¹ and the USA, respectively.² With approximately 370 000 new cases each year in Europe¹ and 210 000 in the USA,³ breast cancer represents a major public health concern and a leading cause of death among women. Identification and measurement of molecular markers that are predictive of response to therapy enable more selective and

effective utilization of treatments, and should lead to overall improvement in patient survival rates.

Approximately 15–25% of patients with breast cancer have tumors that overexpress human epidermal growth factor receptor 2 (HER2).^{4,5} HER2-positive breast cancer is associated with aggressive tumor growth and poor prognosis, especially in patients with node-positive disease.^{6–9} Evidence suggests that HER2 status may be a predictor of response to chemotherapy and hormonal therapy in breast cancer patients^{10–12} and an essential predictor of response to the monoclonal anti-HER2 antibody trastuzumab (Herceptin®).

Significant clinical benefit with trastuzumab treatment has been demonstrated in HER2-positive metastatic breast cancer^{13–16} and, more recently, in

Correspondence: Professor M Dowsett, PhD, Academic Department of Biochemistry, Royal Marsden Hospital, London SW3 6JJ, UK.

E-mail: mitch.dowsett@icr.ac.uk

Received 12 October 2006; revised 12 February 2007; accepted 14 February 2007; published online 30 March 2007

patients with early breast cancer, where reductions in the risk of relapse of approximately 50% have been reported.^{17–19} Consequently, practice guidelines now recommend that HER2 status should be evaluated in all primary breast cancer patients at diagnosis so that optimal patient management can be provided.^{9,20–22} Strict standardization of HER2 testing is necessary to achieve accurate HER2 status determination, identifying those who will gain the greatest benefit from trastuzumab and avoiding unnecessary treatment of patients who are unlikely to respond.

Current HER2-testing systems

Overexpression of HER2 is most commonly caused by amplification of the *HER2* gene,^{8,23} which results in increased HER2 mRNA levels and concomitant overexpression of the HER2 receptor on the tumor cell surface. There is no ‘gold standard’ for HER2 testing, but immunohistochemistry and fluorescence *in situ* hybridization (FISH) are the most commonly used techniques. In addition, chromogenic *in situ* hybridization (CISH) has recently been validated as an alternative to FISH.^{24,25}

Immunohistochemistry uses antibodies to detect expression of HER2 protein on the surface of tumor cells. The level of HER2 protein expression is assessed semi-quantitatively by the intensity and percentage of staining, and scored on a scale of 0–3+ where scores of 0 and 1+ are categorized as negative, 2+ as equivocal, and 3+ as positive.

FISH and CISH are based on the determination of *HER2* gene copy number and use DNA probes. With FISH, fluorescently labeled probes for both *HER2*-specific DNA sequences and the centromere of chromosome 17 (CEP17) are frequently used. The *HER2* fluorescent signal is usually expressed as a ratio relative to the signal for CEP17. CISH, although not used in this study, is an emerging alternative to FISH that uses a peroxidase-labeled probe with chromogenic detection, rather than fluorescent dye, to detect the *HER2* gene. This has the advantage that staining remains stable for a longer period and can be quantified with a standard light microscope. CISH results are based only on *HER2* gene copy number. Control for chromosome 17 copy number in borderline cases requires staining of another sequential slide.

As the results obtained with FISH and CISH are numeric, these tests are more objective and quantitative than immunohistochemistry. Nevertheless, high levels of concordance (90–100%) have been reported between FISH, CISH, and immunohistochemistry.^{24–28}

HER2-testing algorithm

Employing immunohistochemistry as the first-line testing method allows identification of HER2-positive

patients (3+) who may benefit from trastuzumab therapy, whereas HER2-negative patients (0/1+) can be excluded. A proportion of specimens defined as equivocal by immunohistochemistry (2+) must be retested by FISH or CISH to determine HER2 status. This procedure will ensure that all patients who may benefit from trastuzumab are identified (Figure 1). First-line testing can also be performed by FISH or CISH, as shown in Figure 1.²⁵

As current HER2 tests are subject to both analytical and interobserver variation, validation by laboratory proficiency testing is important to improve standardization. Although some quality assessment initiatives already exist, for example, the United Kingdom National External Quality Assessment Service (UK NEQAS), there is a need for more programs to ensure a high standard of validation for HER2-testing methodology.

The objective of this study was to assess immunohistochemistry and FISH interlaboratory consensus between five highly experienced international pathology testing centers with a range of breast cancer specimens and to identify factors that may contribute to discordant results. The study also aimed to evaluate use of a slide-exchange program as a quality assessment instrument.

Materials and methods

Study Design

A slide-exchange program was used to compare immunohistochemistry and FISH testing results between five pathology reference centers in The Netherlands, Canada, France, Belgium, and Germany.

The study included five testing rounds at approximately 2-month intervals. In each round, immunohistochemistry and FISH testing were performed on separate sets of invasive breast cancer specimens.

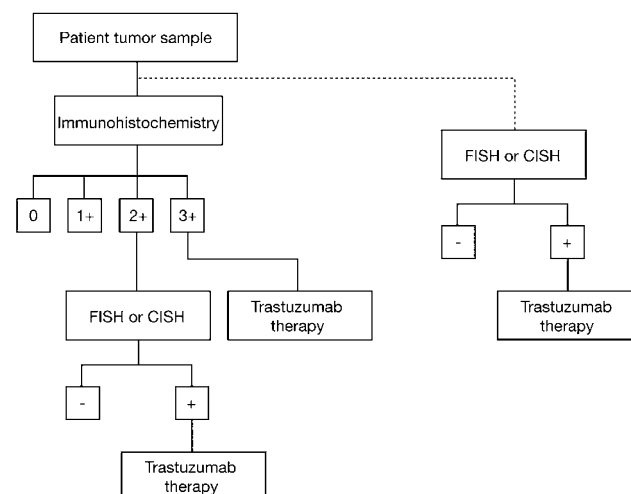


Figure 1 HER2-testing algorithm.²⁵

Thus, a total of 20 immunohistochemistry and 20 FISH breast cancer specimens were evaluated by each of the testing centers over the course of the study. The study was coordinated and the results analyzed by an independent coordinator (Professor Mitch Dowsett, UK), who had no relationship with, or role at, any of the reference centers.

Specimen Selection and Sending of Samples

Each of the five testing centers was designated in turn to select and dispatch the invasive breast cancer specimens to the other four centers. In each testing round, two specimen sets (A and B) of four different invasive breast tumors (ie a total of eight different specimens), which had been previously tested for HER2 status by immunohistochemistry and FISH, respectively, were selected by the sending center. Selected specimens were requested by the coordinator such that they would be representative of a range of HER2 immunohistochemistry expression or FISH amplification levels. The specimens were deliberately selected to include a relatively high proportion of equivocal cases.

All breast cancer specimens were from routine diagnostic practice and had been fixed with formalin (12–48 h) and embedded into paraffin blocks. Tissue sections (4–6 μ m thick) were mounted onto silane-coated slides. Fifteen slides were prepared from each of the eight tumor specimens. Three slides from each specimen were sent to each of the other four testing centers in a blinded manner; the three remaining slides from each specimen were retained by the sending laboratory for its own evaluation (Figure 2).

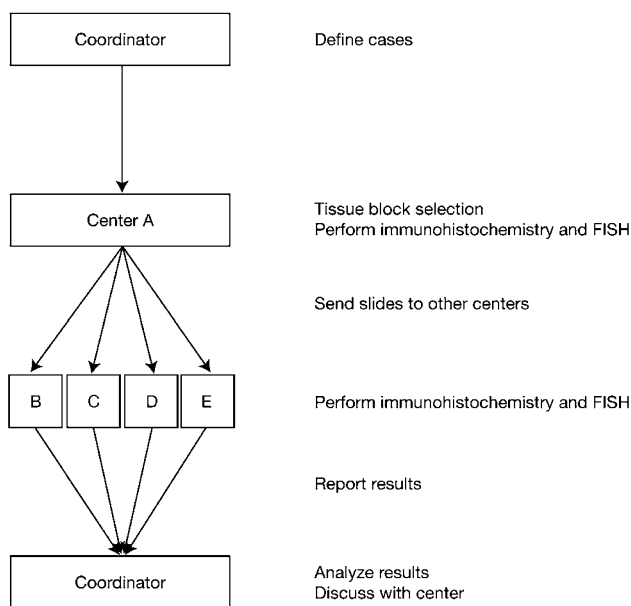


Figure 2 Example of workflow in a testing round, where specimens were selected and sent by center A to centers B, C, D, and E.

Specimen Analysis

Set A: analysis of immunohistochemistry concordance

Each testing center, including the sending center, analyzed the HER2 status of set A specimens by immunohistochemistry using the HercepTest™ test kit (DAKO, Glostrup, Denmark), according to the manufacturer's instructions and recommendations for scoring (in centers B, C, and D, the HercepTest™ kit was not in routine use). Appropriate control specimens were also tested. Immunohistochemistry specimens were scored as 0 (negative), 1+ (negative), 2+ (equivocal), or 3+ (positive) according to the HercepTest™ kit instructions. Specimens scored as equivocal (2+) by any center were subsequently retested by all centers using FISH.

Set B: analysis of FISH concordance

FISH analysis of set B specimens was carried out by each testing center, including the sending center, using the PathVysion™ kit (Vysis/Abbott, IL, USA), according to the manufacturer's instructions and recommendations for scoring. FISH scores were based on the ratio of *HER2:CEP17* signals and were categorized as negative (ratio <2) or positive (ratio ≥ 2.0).

Results from each testing round were sent to the independent coordinator. A final analysis of the results was conducted by the coordinator after completion of all five testing rounds. Consensus among the testing centers for each of the HER2-testing techniques was defined as the percentage of centers with the modal score for each immunohistochemistry or FISH specimen tested.

Results

Analysis of Immunohistochemistry Concordance

The results of the immunohistochemistry analysis of 20 invasive breast cancer specimens are presented in Table 1. Complete consensus between the centers was achieved for nine of the 20 immunohistochemistry specimens. Differences between laboratories were observed with respect to equivocal results (2+). For eight immunohistochemistry specimens, there was at least one center that reported negative HER2 status (0/1+), while others reported equivocal HER2 status. For a further three specimens, at least one center reported positive status (3+) while others reported equivocal status. There was no discordance between the five testing centers for any of the specimens at the level of diagnostic decision, that is, no specimen was categorized as positive at one or more centers but negative at other centers.

In line with recommendations in the HER2-testing algorithm, specimens scored as 2+ using immunohistochemistry at any testing center were retested at all centers using FISH. Two of the 11 specimens (A1

and A3) scored as 2+ using immunohistochemistry were unavailable for retesting with FISH. Mean FISH *HER2*:*CEP17* ratios for the nine retested specimens are shown in Table 2. Of these specimens, five were categorized as negative by FISH in all centers (A6, A11, A12, A14, and A19) and one (A15) was categorized as positive in all centers. Samples A5 and A17 were scored as negative in four centers and positive in one center, while sample A8 was scored as positive in three centers and negative in two.

Table 3 shows the final categorization for all 20 immunohistochemistry specimens after initial immunohistochemistry testing and FISH retesting. After FISH retesting, complete consensus between the five testing centers was achieved for 15 of 18

specimens (83%; two specimens were unavailable for retesting) (Table 3). FISH testing resulted in diagnostic discordance (ie positive vs negative categorization) between the participating centers for specimens A5, A8, and A17.

There were 32 results categorized as equivocal by immunohistochemistry, of which seven related to the two specimens not available for FISH retesting. Of the remaining 25 results categorized as equivocal by immunohistochemistry, 20 (80%) were categorized as negative on FISH retesting and four (16%) were categorized as positive (one specimen produced no signal). In three cases, FISH retesting of specimens that had previously been categorized by a center as positive or negative by immunohistochemistry produced conflicting diagnostic results: center D categorized specimen A5 as positive by immunohistochemistry (3+) but negative with FISH (*HER2*:*CEP17* ratio = 1.7), while centers A and C categorized specimen A8 as negative by immunohistochemistry (0/1+) but positive by FISH (*HER2*:*CEP17* ratios = 2.7 and 2.5, respectively).

Table 1 Analysis of immunohistochemistry concordance: categorization of specimens and consensus among testing centers

Specimen	Center					Consensus (%)
	A	B	C	D	E	
A1	E	E	N	E	N	60
A2	N	N	N	N	N	100
A3	P	E	E	E	E	80
A4	P	P	P	P	P	100
A5	E	E	E	P	E	80
A6	N	E	N	E	N	60
A7	P	P	P	P	P	100
A8	N	E	N	E	E	60
A9	P	P	P	P	P	100
A10	N	N	N	N	N	100
A11	E	E	E	E	N	80
A12	E	E	N	E	N	60
A13	N	N	N	N	N	100
A14	E	E	N	E	E	80
A15	P	P	E	P	P	80
A16	P	P	P	P	P	100
A17	N	E	N	N	N	80
A18	P	P	P	P	P	100
A19	N	E	E	E	N	60
A20	N	N	N	N	N	100

N: negative (0 or 1+); E: equivocal (2+); P: positive (3+).

Table 2 Analysis of equivocal immunohistochemistry specimens using FISH: mean *HER2*:*CEP17* ratios

Specimen	Center				
	A	B	C	D	E
A5	1.1	1.48	2.15	1.7	1.31
A6	0.9	1.2	1.3	1.0	1.23
A8	2.7	1.73	2.5	1.8	2.91
A11	No signal	1.31	1.3	1.4	1.33
A12	1.2	1.14	1.2	1.2	1.47
A14	1.2	1.05	1.5	1	1.26
A15	Amplified	4.37	>4	4.5	5.51
A17	1.3	2.1	1.7	1.09	1.42
A19	1.2	1.72	1.3	1.65	1.0

Analysis of FISH Concordance

Complete concordance between all five testing centers was found for 16 of 20 specimens analyzed by FISH (80%; six negative, 10 positive) (Table 4).

All four discordant FISH specimens were scored as having *HER2*:*CEP17* ratios within the range 1.7–2.3 by at least one center. The four specimens for which the centers did not agree had mean (range) *HER2*:*CEP17* ratios of 2.00 (0.92–2.70), 1.48 (1.10–2.00), 1.72 (1.00–2.92), and 1.82 (1.23–2.61).

Among the different testing centers, FISH *HER2*:*CEP17* ratios were highest from center D for 10 of the 20 specimens; by contrast, centers B and C each reported the highest score for just one specimen. Mean ratios from across the five testing centers were calculable for 12 of the specimens. The mean differences for each center from the group mean were –0.07, –0.11, –0.14, 0.41, and –0.08 for centers A, B, C, D, and E, respectively. The corresponding median differences from the group mean were –0.02, –0.13, –0.13, 0.17, and –0.06, respectively.

Discussion

This slide-exchange ring study shows that under standardized conditions, there is a high level of consensus between pathology testing centers for *HER2* testing by both immunohistochemistry and FISH. It also highlights that some discordance occurs, predominantly for borderline-positive samples, even between expert laboratories. It should be noted when considering the level of discordance reported in this study that specimens were pre-selected to contain a higher proportion of equivocal cases than would be expected in the general

Table 3 Analysis of immunohistochemistry specimens, including retesting of equivocal specimens by FISH: recategorization of specimens and consensus among testing centers

Specimen	Center					Consensus (%)
	A	B	C	D	E	
A1 ^a	E	E	N	E	N	60
A2	N	N	N	N	N	100
A3 ^a	P	E	E	E	E	80
A4	P	P	P	P	P	100
A5	E (FISH N)	E (FISH N)	E (FISH P)	P (FISH N)	E (FISH N)	Remained at 80
A6	N (FISH N)	E (FISH N)	N (FISH N)	E (FISH N)	N (FISH N)	Increased to 100
A7	P	P	P	P	P	100
A8	N (FISH P)	E (FISH N)	N (FISH P)	E (FISH N)	E (FISH P)	Remained at 60
A9	P	P	P	P	P	100
A10	N	N	N	N	N	100
A11	E (no signal)	E (FISH N)	E (FISH N)	E (FISH N)	N (FISH N)	Increased to 100
A12	E (FISH N)	E (FISH N)	N (FISH N)	E (FISH N)	N (FISH N)	Increased to 100
A13	N	N	N	N	N	100
A14	E (FISH N)	E (FISH N)	N (FISH N)	E (FISH N)	E (FISH N)	Increased to 100
A15	P (FISH P)	P (FISH P)	E (FISH P)	P (FISH P)	P (FISH P)	Increased to 100
A16	P	P	P	P	P	100
A17	N (FISH N)	E (FISH P)	N (FISH N)	N (FISH N)	N (FISH N)	Remained at 80
A18	P	P	P	P	P	100
A19	N (FISH N)	E (FISH N)	E (FISH N)	E (FISH N)	N (FISH N)	Increased to 100
A20	N	N	N	N	N	100

P: positive (*HER2*:CEP17 ratio ≥ 2); N: negative (*HER2*:CEP17 ratio < 2).

^aA1 and A3 were unavailable for FISH retesting.

Bold type indicates samples with discordant results.

population: 32 of the 100 immunohistochemistry results were rated as equivocal compared with about 15% in routine practice.^{29,30} Under routine conditions with fewer equivocal cases, an even lower level of discordance than that reported in this analysis might be expected.

The results presented here illustrate the difficulty, even for experienced laboratories, in determining the *HER2* status of equivocal cases. Although there were no cases where, by immunohistochemistry testing alone, the same immunohistochemistry specimen was categorized as positive by some laboratories and negative by others, in $>50\%$ of cases, specimens were categorized as equivocal (2+) by one or more centers, while others gave a clear positive or negative categorization. Centers B and D categorized twice as many specimens as equivocal than center E.

The nature of this quality assessment study did not allow variability between the centers to be ascribed with certainty to differences in the product of the immunohistochemistry, as opposed to differences in the interpretation of the product. Inter-observer variability is the most likely explanation for the differences observed. However, centers B and D did not routinely use the HercepTest™ in their everyday practice, so it is possible that their use of it may have varied from the others in detail and contributed to their having the highest number of equivocal scores. It is possible that application of image analysis might have improved immunohistochemistry concordance between the centers.³¹

On retesting of equivocal cases using FISH, the concordance rate increased to 15 of 18 specimens (83%). Of the 25 equivocal immunohistochemistry categorizations retested by FISH, 80% were recategorized as negative and 16% as positive, which is within the range reported when equivocal specimens are retested using FISH in routine practice. These data illustrate the importance of retesting equivocal specimens as specified by the *HER2*-testing algorithm,³² since, based on the above results, one in six patients who could benefit from trastuzumab may have been excluded from therapy.

Considering the three immunohistochemistry specimens, where complete concordance between centers was not achieved even with FISH retesting, in one case (A8) two centers (centers B and D) reported *HER2*:CEP17 ratios of 1.73 and 1.8, while the others reported ratios of >2 . In the two remaining cases (A5 and A17), one center reported FISH *HER2*:CEP17 ratios slightly >2 (A5 center C *HER2*:CEP17 = 2.15; A17 center B *HER2*:CEP17 = 2.1), while all other centers reported ratios of <2 . Detailed retrospective assessment showed that A17 was a highly heterogeneous specimen, with small foci of 2+ staining with immunohistochemistry (comprising $<10\%$ of the total cells). By focusing only on these areas, center B recorded a *HER2*:CEP17 ratio of 2.1. This was in contrast to the other four centers, in which both positive and negative cells were counted, resulting in *HER2*:CEP17 ratios of 1.09–1.7.

Table 4 Analysis of FISH concordance: mean *HER2*:CEP17 ratios, categorization of specimens, and consensus among testing centers

Specimen	Center					Consensus (%)	Mean score ^a
	A	B	C	D	E		
B1	N	N	N	N	N	100	
	1.00	1.03	1.10	1.31	1.08		1.10
B2	N	N	N	N	N	100	
	1.00	1.74	1.00	1.08	1.07		1.18
B3	P	P	P	P	P	100	
	8.90	>4	>10	6.24	>10		—
B4	N	P	P	P	P	80	
	0.92	2.17	2.70	2.07	2.15		2.00
B5	P	P	P	P	P	100	
	3.49	3.48	3.20	5.91	3.47		3.91
B6	N	N	N	N	N	100	
	1.16	1.04	1.20	1.28	1.11		1.16
B7	P	P	P	P	P	100	
	6.28	>4 ^b	>8	5.63	>10		—
B8	P	P	P	P	P	100	
	4.15	3.3 ^c	3.9	4.01	3.8		—
B9	N	N	N	N	N	100	
	1.12	0.95	1.1	1.17	1.10		1.09
B10	P	P	P	P	P	100	
	2.41	2.6	2.6	3.22	>6.4		—
B11	P	P	P	P	P	100	
	5.83	>4	>10	>10	>9.2		—
B12	P	P	P	P	P	100	
	3.03	2.98	3	3.6	2.96		3.11
B13	N	N	N	N	N	100	
	1.15	1.00	1	1.5	1.07		1.14
B14	P	P	P	— ^d	P	100	
	12.2	>4	>5		9		—
B15	P	P	P	P	P	100	
	2.21	>2	>2	3	2.4		—
B16	P	N	N	N	N	80	
	2.00	1.49	1.2	1.1	1.6		1.48
B17	N	N	N	P	N	80	
	1.74	1.35	1	2.92	1.6		1.72
B18	P	P	P	P	P	100	
	4.06	>4	>5	6.3	>5		—
B19	N	N	N	N	N	100	
	1.03	1	1	1.09	1.19		1.06
B20	P	N	N	P	N	60	
	2.24	1.23	1.59	2.61	1.42		1.82

N: negative (*HER2*:CEP17 ratio <2); P: positive (*HER2*:CEP17 ratio ≥2).

^a*HER2*:CEP17 ratios, where calculable.

^bIn clusters.

^cIn some clusters.

^dNo invasive tumor.

Bold type indicates samples with discordant results.

These results highlight a difficulty in interpreting borderline FISH scores and in assessing cases with intratumoral heterogeneity. It has been suggested that such cases may comprise approximately 1% of all breast tumors.³³ Whether to assess the majority of cells in a specimen or to focus only on foci of positively stained cells remains a matter of debate and, until more treatment–response data are available, it is difficult to determine the clinical implications of trastuzumab treatment in these patients.

The analysis of FISH concordance between the testing laboratories revealed a concordance for 16 of

the 20 specimens (80%). All four discordant FISH specimens were scored as having *HER2*:CEP17 ratios within the range 1.7–2.3 by at least one center. In one case (B16), the lack of concordance was due to one center (Center A) reporting a ratio of 2.0, which was classified as positive, whereas the other centers reported ratios of <2.0 (negative). None of these discordant specimens was retested. It is perhaps inevitable that some discordance will be encountered around borderline levels for positive or negative scores and in these cases it cannot be definitively stated which result is ‘correct’ and which is ‘incorrect’. It has recently been suggested that borderline cases may even constitute a unique tumor type, with implications for treatment response.³⁴

Overall, the results of this study support the *HER2*-testing algorithm, which is adequate for the vast majority of specimens. However, for specimens that fail to be resolved by the first round of FISH analysis, it is recommended that FISH or CISH retesting should be considered.^{25,32} The discrepancies observed with the FISH analysis highlight that the exclusive use of FISH for *HER2* testing could lead to misdiagnosis in some cases. It is important to note that by using both immunohistochemistry and FISH, as recommended in the *HER2*-testing algorithm, the chances of misdiagnosis are reduced.

The quantitative nature of FISH analysis makes interobserver variability much less of an issue than with immunohistochemistry. Nevertheless, one center consistently scored specimens higher or lower than other centers in many cases: 10 of 12 specimens from center D were higher than the overall group mean. Although immunohistochemistry and FISH were analyzed by all centers using standardized procedures, small variations in sample processing and the relative experience of laboratory personnel could potentially influence results obtained by different centers. The use of validated in-house immunohistochemistry protocols that differed slightly from the HercepTest™ may also have contributed to the level of discordance.

A high standard of validation for any *HER2*-testing methodology is needed for optimal identification of patients likely to respond to trastuzumab therapy.³⁵ In this study, samples were preselected, testing reagents were prescribed, and experienced personnel performed the *HER2* testing, yet there was still discordance. This highlights the inherent difficulties encountered during *HER2* testing using immunohistochemistry and FISH, even for laboratories with extensive experience of these procedures. Consequently, it is expected that inexperienced laboratories will have greater problems interpreting *HER2* status results. This study emphasizes the need for rigorous quality-control procedures for the preparation and analysis of specimens and the validation of results from less experienced laboratories by a centralized reference laboratory. Organizations such as UK NEQAS have a

role to play in ensuring a high standard of quality assessment; however, at present, UK NEQAS only assesses methodologies and not the interpretation of results. By adopting a slide-exchange program such as that used in the current ring study, even laboratories with considerable experience may identify not only technical issues but also discrepancies in the interpretation of HER2 testing, which may be remedied. In addition to national schemes, reference laboratories should consider taking a leading role in the initiation of such quality-control studies, as newly established laboratories are likely to benefit from their experience.

Acknowledgements

Frederique Penault-Llorca thanks Anne Cayre for helpful advice and co-interpretation of FISH specimens, and Matthieu Roche for technical assistance. Wedad Hanna would like to acknowledge the technical assistance of Kevin Kwok and the clerical assistance of Karen Laws.

References

- Boyle P, Ferlay J. Cancer incidence and mortality in Europe, 2004. *Ann Oncol* 2005;16:481–488.
- Ries LAG, Harkins D, Krapcho M, *et al*. National Cancer Institute. SEER Cancer Statistics Review, 1975–2003. 2006. Available at: http://seer.cancer.gov/csr/1975_2003/.
- American Cancer Society. Breast cancer facts & figures 2005–2006. 2006. Available at: <http://www.cancer.org/downloads/STT/CAFF2005BrF.pdf>.
- Hynes NE, Stern DF. The biology of erbB-2/neu/HER-2 and its role in cancer. *Biochim Biophys Acta* 1994; 1198:165–184.
- Owens MA, Horten BC, Da Silva MM. HER2 amplification ratios by fluorescence *in situ* hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues. *Clin Breast Cancer* 2004; 5:63–69.
- Press MF, Bernstein L, Thomas PA, *et al*. HER-2/neu gene amplification characterized by fluorescence *in situ* hybridization: poor prognosis in node-negative breast carcinomas. *J Clin Oncol* 1997;15:2894–2904.
- Ménard S, Fortis S, Castiglioni F, *et al*. HER2 as a prognostic factor in breast cancer. *Oncology* 2001; 61(Suppl 2):67–72.
- Ross JS, Fletcher JA, Bloom KJ, *et al*. Targeted therapy in breast cancer. The HER-2/neu gene and protein. *Mol Cell Proteomics* 2004;3:379–398.
- Goldhirsch A, Glick JH, Gelber RD, *et al*. Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. *Ann Oncol* 2005;16:1569–1583.
- Hanna W. Testing for HER2 status. *Oncology* 2001; 61(Suppl 2):22–30.
- Perez EA. HER-2 as a prognostic, predictive, and therapeutic target in breast cancer. *Cancer Control* 1999;6:233–240.
- Paik S, Park C. HER-2 and choice of adjuvant chemotherapy in breast cancer. *Semin Oncol* 2001;28: 332–335.
- Slamon DJ, Leyland-Jones B, Shak S, *et al*. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001;344:783–792.
- Vogel CL, Cobleigh MA, Tripathy D, *et al*. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;20:719–726.
- Baselga J, Carbonell X, Castañeda-Soto N-J, *et al*. Phase II study of efficacy, safety, and pharmacokinetics of trastuzumab monotherapy administered on a 3-weekly schedule. *J Clin Oncol* 2005;23:2162–2171.
- Marty M, Cognetti F, Maraninchi D, *et al*. Randomized phase II trial of the efficacy and safety of trastuzumab combined with docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer administered as first-line treatment: the M77001 study group. *J Clin Oncol* 2005;23:4265–4274.
- Slamon D, Eiermann W, Robert N, *et al*. Phase III randomized trial comparing doxorubicin and cyclophosphamide followed by docetaxel (AC→T) with doxorubicin and cyclophosphamide followed by docetaxel and trastuzumab (AC→TH) with docetaxel, carboplatin and trastuzumab (TCH) in HER2 positive early breast cancer patients: BCIRG 006 study [abstract]. *Breast Cancer Res Treat* 2005;94(Suppl 1):S5.
- Piccari-Gebhart MJ, Procter M, Leyland-Jones B, *et al*. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005;353: 1659–1672.
- Romond EH, Perez EA, Bryant J, *et al*. Doxorubicin and cyclophosphamide followed by paclitaxel with or without trastuzumab as adjuvant therapy for patients with HER-2 positive operable breast cancer. Combined analysis of NSABP-B31/NCCTG-N9831. Slide presentation at 41st ASCO Annual Meeting, Orlando, Florida, USA, May 13–17, 2005. Available at: http://www.asco.org/ac/1,1003,12-002511-00_18-0034-00_19-005817-00_21-001,00.asp.
- Bast Jr RC, Ravdin P, Hayes DF, *et al*. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. *J Clin Oncol* 2001;19:1865–1878.
- Carlson RW, McCormick B. Update: NCCN breast cancer clinical practice guidelines. *J Natl Compr Canc Netw* 2005;3(Suppl 1):S7–S11.
- National Comprehensive Cancer Network. Clinical Practice Guidelines in Oncology—v.2.2006. Breast Cancer. Version 2.2006. 2006. Available at: http://www.nccn.org/professionals/physician_gls/PDF/breast.pdf.
- Hicks DG, Tubbs RR. Assessment of the HER2 status in breast cancer by fluorescence *in situ* hybridization: a technical review with interpretive guidelines. *Hum Pathol* 2005;36:250–261.
- van de Vijver MJ, Bilous M, Hanna W, *et al*. Correlation of chromogenic *in-situ* hybridisation (CISH) with FISH and IHC for assessment of HER2 gene amplification: an international validation ring study [abstract]. *EJC Suppl* 2004;2:100.
- Hanna WM, Kwok K. Chromogenic *in-situ* hybridization: a viable alternative to fluorescence *in-situ*

- hybridization in the HER2 testing algorithm. *Mod Pathol* 2006;19:481–487.
- 26 Dandachi N, Dietze O, Hauser-Kronberger C. Chromogenic *in situ* hybridization: a novel approach to a practical and sensitive method for the detection of HER2 oncogene in archival human breast carcinoma. *Lab Invest* 2002;82:1007–1014.
- 27 Arnould L, Denoux Y, MacGrogan G, *et al*. Agreement between chromogenic *in situ* hybridisation (CISH) and FISH in the determination of HER2 status in breast cancer. *Br J Cancer* 2003;88:1587–1591.
- 28 Bilous M, Morey A, Armes J, *et al*. The inter-laboratory reproducibility of CISH testing for HER2 and correlation with IHC and FISH results. Poster presented at the EBCC, Hamburg, Germany, 16–20 March 2004.
- 29 Sawaki M, Ito Y, Akiyama F, *et al*. High prevalence of HER-2/*neu* and p53 overexpression in inflammatory breast cancer. *Breast Cancer* 2006;13:172–178.
- 30 Dowsett M, Bartlett J, Ellis IO, *et al*. Correlation between immunohistochemistry (HercepTest) and fluorescence *in situ* hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *J Pathol* 2003;199:418–423.
- 31 Wang S, Saboorian MH, Frenkel EP, *et al*. Assessment of HER-2/*neu* status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence *in situ* hybridization assay as the standard. *Am J Clin Pathol* 2001;116:495–503.
- 32 Bilous M, Dowsett M, Hanna W, *et al*. Current perspectives on HER2 testing: a review of national testing guidelines. *Mod Pathol* 2003;16:173–182.
- 33 Hanna W, Nofech-Mozes S, Kahn H. Intratumoral heterogeneity of HER2/*neu* in breast cancer—a rare event or an overlooked phenomenon? Abstract no. 123 presented at USCAP 2006. Available at: http://www.abstracts2view.com/uscap06/view.php?nu=USCAP06L_123.
- 34 Killeen JL, Ortega-Lopez A, Shaha J, *et al*. Pathology of borderline HER-2/*neu* breast carcinoma: a biologically distinct phenotype. *Breast Cancer Res Treat* 2006;98:99–108.
- 35 Carlson RW, Moench SJ, Hammond MEH, *et al*. HER2 testing in breast cancer: NCCN Task Force report and recommendations. *J Natl Compr Canc Netw* 2006; 4(Suppl 3):S1–S22.