EUROPEAN JOURNAL OF CANCER 42 (2006) 2946–2953

available at www.sciencedirect.com

journal homepage: www.ejconline.com

High inter-observer agreement in immunohistochemical evaluation of HER-2/*neu* expression in breast cancer: A multicentre GEFPICS study ☆

Magali Lacroix-Triki[a,*], *Simone Mathoulin-Pelissier*[b], *Jean-Pierre Ghnassia*[c], *Gaetan Macgrogan*[d], *Anne Vincent-Salomon*[e], *Véronique Brouste*[b], *Marie-Christine Mathieu*[f], *Pascal Roger*[g], *Frédéric Bibeau*[h], *Jocelyne Jacquemier*[i], *Frédérique Penault-Llorca*[j], *Laurent Arnould*[k]

[a]*Department of Pathology, Institut Claudius Regaud, 20–24 rue du pont Saint Pierre, 31052 Toulouse Cedex, France*
[b]*Department of Biostatistics, Institut Bergonié, Bordeaux, France*
[c]*Centre Paul Strauss, Strasbourg, France*
[d]*Institut Bergonié, Bordeaux, France*
[e]*Institut Curie, Paris, France*
[f]*Institut G. Roussy, Villejuif, France*
[g]*CHRU Montpellier, France*
[h]*Centre Val d'Aurelle, Montpellier, France*
[i]*Institut P. Calmette, Marseille, France*
[j]*Centre J. Perrin, Clermont-Ferrand, France*
[k]*Centre G.F. Leclerc, Dijon, France*

ARTICLE INFO

Article history:
Received 26 January 2006
Received in revised form
31 March 2006
Accepted 26 June 2006
Available online 20 September 2006

Keywords:
Breast cancer
HER-2/*neu*
Immunohistochemistry
Inter-observer reproducibility
Quality assurance
Ring study

ABSTRACT

The accurate determination of HER-2 in invasive breast cancer has become a critical issue, particularly in the context of the results of recent trastuzumab (Herceptin®) adjuvant trials. This multicentre study evaluated inter-observer reproducibility in interpretation of HER-2 immunostains performed in different laboratories according to their in-house technique. A total of 74 HER-2 immunostains were evaluated by 16 pathologists and by a central review committee. As determined by central review, the HER-2 score was 0 in 33 cases (44%), 1+ in 10 cases (13%), 2+ in 9 cases (12%) and 3+ in 23 cases (31%). The overall kappa value was good (kappa = 0.75). Agreement was excellent for the 0/1+ group (kappa = 0.85) and for the 3+ group (kappa = 0.82). As expected, the score 2+ group showed poor agreement (kappa = 0.38). A quality assurance process showed that ring studies and adherence to national guidelines greatly improve inter-observer reproducibility.

☆ GEFPICS: Groupe d'Etude des Facteurs Pronostiques par Immunohistochimie dans le Cancer du Sein (Group for the Evaluation of Immunohistochemical Prognostic Factors in Breast Cancer). Members of the group: the authors and M. Antoine, A. Balaton, M.-C. Baranzelli, V. Becette, J.P. Bellocq, M.-C. Blanc-Fournier, Y. Denoux, F. Ettore, M. Frassard, V. Fridman, B. Lannes, C. Migeon, C. Rigaud, C. Sagan, B. Sigal-Zafrani, I. Treilleux, V. Verrièle.

* *Corresponding author:* Tel.: +335 61 42 42 31; fax: +335 61 42 46 02.
E-mail address: lacroix.magali@claudiusregaud.fr (M. Lacroix-Triki).

0959-8049/$ - see front matter
doi:10.1016/j.ejca.2006.06.020

1. Introduction

Accurate determination of HER-2 status has become of major clinical importance in breast cancer. HER-2 over-expression, which is observed in 10–30% of breast carcinomas, is associated with poor clinical outcome, as it is correlated with shorter disease-free survival.[1] Moreover, HER-2 over-expression predicts for response to anti-HER-2 therapy with the recombinant humanised anti-p185$^{HER-2/neu}$ antibody trastuzumab (Herceptin®). Promising results of three trastuzumab adjuvant trials were presented at the American Society of Clinical Oncology 2005 annual meeting.[2,3] These three trials: the National Surgical Adjuvant Breast and Bowel Project B-31 (NSABP B-31); the North Central Cancer Treatment Group N9831 (NCCTG N9831); and the HERceptin® Adjuvant (HERA) trial, showed a 52% risk reduction for recurrence with adjuvant trastuzumab compared with controls.[2,3] Furthermore, accumulating evidence suggests that HER-2 over-expression may also predict sensitivity to anthracycline-based chemotherapy [4–6] and may help in the choice of endocrine therapy.[7]

Currently, no single assay is globally accepted as the gold standard for HER-2 determination, but among the three validated techniques, i.e. immunohistochemistry (IHC), fluorescence *in situ* hybridisation (FISH) and chromogenic *in situ* hybridisation (CISH), IHC is the most commonly used technique and the only one to study the treatment's target, i.e. the HER-2 protein. Indeed, IHC is a reliable, easy to perform and accessible technique, which is far less expensive and time-consuming than FISH. Moreover, several studies have demonstrated a high concordance between results for IHC and FISH.[8–12] However, the choice of technique for HER-2 determination remains a matter of debate. Firstly, the IHC technique was criticised because of a lack of inter-laboratory reproducibility due to the variability in fixation, tissue processing, IHC protocol, anti-HER-2 antibodies and scoring system used in each laboratory, i.e. the so-called 'real world'. Secondly, HER-2–IHC was considered to be a subjective test for the assessment of staining intensity and percentage of labelled tumoural cells, and hence prone to inter-observer variability.[13] In a clinical laboratory assay study of HER-2 testing, the College of American Pathologists thus emphasised the importance of adhering to standardised protocols for IHC-HER-2, claiming that the initial validation of assays against a gold standard is mandatory for accuracy.[14]

In this context, the French multicentre group Groupe d'Etude des Facteurs Pronostiques par Immunohistochimie dans le Cancer du Sein (GEFPICS) performed a study comparing the HER-2 status of 119 breast invasive carcinomas, determined by IHC in 12 different laboratories before and after calibration by reference to FISH on the corresponding frozen tissue sections.[15] This study showed that a high accuracy of IHC could be obtained for the determination of HER-2 status in all laboratories using their in-house IHC technique, provided that a calibration process was performed. Therefore, after examining the question of inter-laboratory variability, we aimed to study inter-observer reproducibility, particularly as there are few studies addressing this problem.

Inter-observer reproducibility was determined among 19 pathologists from 19 different institutions, who were asked to interpret the IHC–HER-2 stains of 75 breast invasive carcinomas.

2. Material and methods

2.1. Selection of cases

The cases for this study were derived from a previous French multicentre GEFPICS study, which concerned the calibration of IHC for assessment of HER-2 in breast cancer. Details of this study have been reported elsewhere.[15] Briefly, 12 different French laboratories selected 119 invasive breast carcinomas. Initial criteria for tumour selection were the availability of frozen tissue and previously immunohistochemically determined HER-2 status. Tumours were then secondary selected so as to balance the series with negative and positive cases in a range close to the distribution observed in clinical practice. At the end of this study, the 12 different laboratories had calibrated their HER-2 immunohistochemical technique, by reference to FISH performed on frozen sections. Among this series, 75 cases were used for the present study.

2.2. IHC procedure and FISH analysis

For each case, an IHC–HER-2 stain was performed by the laboratory of origin. Therefore, 4 different fixatives were used in this series (Table 1). Nineteen tumours were fixed in alcohol–formalin–acetic (ethanoic) acid (ethanoic acid 5%/ethanol 100° 75%/water 18%/commercial formalin (methanal) 2%), 43 in neutral buffered formalin, 9 in Hollande–Bouin's fixative and 4 in Bouin's fixative. The monoclonal antibody CB11 (Novocastra, Newcastle, United Kingdom (UK)) was used in 39 cases, the polyclonal antibody A0485 (Dako, Glostrup, Denmark) in 36 cases (Table 1). Details of the IHC procedure for each antibody are described in Table 2. Briefly, 11 out of 12 laboratories used a heat-induced antigen retrieval technique, mainly with a citrate pH 6 buffer. Primary antibody dilution and incubation time varied between the 12 laboratories. All the detection procedures relied on avidin-biotin-based systems, except in one centre which used a dextran polymer enhancing system (Envision®, Dako) for CB11 detection.

The FISH status was known in 72 cases, as it was performed in the previously described study.[15] These FISH experiments were carried out on frozen samples using the

Table 1 – Fixative and antibodies used for the HER-2 immunostains

Fixative	Cases (*n*)	Anti-HER-2	
		CB11[a]	A0485[b]
Acetic (ethanoic) acid /formalin/alcohol	19	15	4
Neutral buffered formalin	43	24	19
Hollande–Bouin's fixative	9	0	9
Bouin's fixative	4	0	4
Total	75	39	36

a Antibody CB11 (Novocastra, Newcastle, United Kingdom (UK)).
b Antibody A0485 (Dako, Glostrup, Denmark).