Concordance of HER2 Immunohistochemistry and Fluorescence In Situ Hybridization Using Tissue Microarray in Breast Cancer

DANIELA FURRER1,2,3, SIMON JACOB4,5,6, CHANTAL CARON4,5, FRANÇOIS SANSCHAGRIN2,4, LOUISE PROVENCHER1,2,4,7 and CAROLINE DIORIO1,2,3,4

1Cancer Research Center at Laval University, and Departments of 3Social and Preventive Medicine, 6Molecular Biology, Medical Biochemistry and Pathology, and 7Surgery, Faculty of Medicine, Laval University, Quebec City, Canada; 2Oncology Axis, CHU of Quebec Research Center, Laval University, Quebec City, Canada; 4Deschênes-Fabia Center for Breast Diseases, and 5Pathology Service, Saint-Sacrement Hospital, Quebec City, Canada

Abstract. Aim: Immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) are common methods for assessment of human epidermal growth factor receptor 2 (HER2) in breast cancer. Materials and Methods: In a cohort of 498 consecutive patients with breast cancer, we examined concordance between IHC and FISH for HER2 on tissue microarray (TMA) sections. In a subset of 116 specimens, we examined HER2 concordance from the block used for diagnostics and a randomly-chosen additional block (a proxy of the core biopsy). Results: Overall concordance between both methods on TMA sections was 93.8% and between HER2, determined on diagnostic and additional blocks, was 93.6% for IHC and 98.0% for FISH. Conclusion: Since some cases were discordant, we suggest that both methods be used for HER2 assessment. The lower concordance rate between diagnostic and additional blocks using IHC compared to FISH suggests a greater variability of IHC staining across tumor regions than for FISH results.

Human epidermal growth factor receptor 2 (HER2) is a transmembrane tyrosine kinase receptor belonging to the family of epidermal growth factor receptors (EGFR) (1). The protein is encoded by the HER2 (ERBB2) gene, which is located on the long arm of chromosome 17 (17q12-21.32) (2). HER2 gene amplification and protein overexpression, which occur in 15 to 20% of patients with breast cancer, are important markers for poor prognosis, including a more aggressive disease and a shorter survival (3). Moreover, HER2-positive status is considered a predictive marker of response to HER2-targeted drugs, including trastuzumab and lapatinib (4). Given its prognostic, predictive and therefore therapeutic implications, an accurate evaluation of HER2 status is crucial for identification of patients who would most likely benefit from targeted anti-HER2 therapies.

Currently, there are several Food and Drug (FDA)-approved methods for the evaluation of HER2 status in breast cancer specimens, including immunohistochemical (IHC) determination of HER2 protein expression or assessment of HER2 gene amplification using in situ hybridization (ISH), most commonly fluorescent ISH (FISH) (5, 6). Patients are eligible for targeted anti-HER2 therapies when their breast cancer specimens overexpress the protein at IHC and are HER2-gene amplified at ISH. Since both IHC and FISH present advantages and disadvantages, there is still no consensus on which method is superior for assessing the HER2 status in breast cancer specimens (7). In 2013, the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) updated the guidelines published in 2007 to clarify the recommendations for HER2 testing in breast cancer specimens (5). In particular, the 2013 ASCO/CAP guidelines include new scoring criteria for IHC and FISH. Moreover, the updated guidelines recommend performing an initial test (IHC or ISH) on core biopsy. If test results are equivocal, reflex testing on tumor specimen section with an alternative assay (IHC or ISH) should be carried out. In addition, repeat-testing should be performed if there is an apparent histopathological discordance with the test result (5). The 2007 ASCO/CAP guidelines recommended to perform HER2 testing on resection specimens and to retest when results were equivocal (6).
Several publications have demonstrated a very good concordance between the results obtained by IHC and FISH for the determination of HER2 status in breast cancer specimens (8, 9). Other studies, however, have reported up to 13% of FISH-positive cases among those scoring negative by IHC (10-15). Considering this discordance, it would therefore be interesting to perform both IHC and FISH on every specimen. Since the vast majority of studies published to date analyzed the concordance between IHC and FISH according to the 2007 ASCO/CAP criteria (6), it is pertinent to analyze the concordance between these two FDA-approved techniques according to the recent ASCO/CAP scoring criteria (5). Moreover, we wished to evaluate the new approach by which HER2 status is determined firstly on core biopsy as recommended by the updated ASCO/CAP guidelines (5).

Tissue microarray (TMA) allows for molecular characterization of large numbers of specimens by means of arranging tissue cores from multiple samples into an empty paraffin block. Large numbers of specimens can therefore be processed under identical conditions and analyzed simultaneously (16, 17). The goal of this study was to analyze the concordance of HER2 status determined by IHC and FISH in 498 consecutive breast cancer specimens on TMA sections and evaluate the impact of the new ASCO/CAP scoring guidelines on the concordance between results. Moreover, in a subset of 116 breast cancer specimens, we aimed to evaluate the concordance between HER2 status from the routine diagnostic paraffin block and a randomly chosen additional block (a proxy of the core biopsy) using TMA.

Materials and Methods

Specimen collection and patient population. The study population has already been described elsewhere (18). Briefly, formalin-fixed, paraffin-embedded breast cancer tissues from 554 consecutive patients with invasive breast carcinoma were used. Patients who received neoadjuvant chemotherapy or with a tumor size smaller than 1 cm on histological slides were excluded from the study. All samples were collected at the Centre des Maladies du Sein Deschênes-Fabia at the Saint-Sacrement Hospital in Québec, Québec, Canada, between February 2011 and April 2012. ASCO/CAP recommendations regarding formalin fixation time and time to fixative were followed (5). Written informed consent was obtained from all participants. Ethical approval of the study was obtained from the Research Ethics Committee of the Centre de Recherche du CHU de Québec ( Permit Number: DR-002-1286).

TMA construction. TMAs were constructed as previously described (18). Briefly, for the 554 consecutive breast cancer specimens, the most representative tissue block from each case was selected by the pathologist for HER2 assessment by IHC and FISH. Two tumoral regions showing the strongest IHC staining were delineated on the IHC slide by the pathologist. Four 0.6 mm tissue cylinders were punched using a manual arraying instrument within these delineated regions (Beecher Instruments, Silver Spring, MD, USA) and were inserted into empty recipient paraffin blocks. We call these TMA blocks “diagnostic TMA”. Among the 554 cases present on the diagnostic TMA, cases that had at least two additional paraffin blocks showing the same histological type as the paraffin block that was used for the construction of the diagnostic TMA were selected. Among these cases, 100 cases were randomly chosen. Suitable blocks were identified using H&E section and two tumoral regions were delineated without previous IHC staining on the H&E section by two trained technologists and verified by a pathologist. Four 0.6 mm tissue cylinders were then punched from these regions. Since we also performed an oversampling of HER2-positive cases (16 cases that were scored as HER2-positive on diagnostic TMA section but that had not been randomly selected were added to the 100 cases), a total of 116 cases were inserted into an empty paraffin block. We call these TMA blocks “random TMA”. Consecutive 4-μm sections were processed by IHC and FISH. From each TMA block one section was stained with H&E for reference histology.

Immunohistochemistry. HER2 protein expression was performed using the FDA-approved HercepTest™ kit (DAKO Diagnostics, Glostrup, Denmark) on an automated immunostaining system (Autostainer; DAKO), according to the manufacturer’s instructions.

FISH. HER2 gene copy number was evaluated using the FDA-approved PathVysion™ HER2 DNA Probe kit (Abbott Molecular, Des Plaines, IL, USA/Inter Medico, Markham, Canada), according to the manufacturer’s instructions.

HER2 evaluation on TMA. A core was considered satisfactory for analysis if tumor tissue occupied >10% of the core area.

Immunohistochemical staining was analyzed visually. Cytoplasmic staining was ignored and only invasive tumor was scored.

Fluorescent signals were evaluated as previously described (18). Briefly, fluorescent signals were evaluated with an epifluorescence microscope (Axio Imager M1; Zeiss, Göttingen, Germany). Automated analysis of fluorescence signals was carried out using the FDA-approved MetaSystems™ image analysis system (19). For cores with equivocal results, automated image analysis was followed by manual counting in at least 60 nonoverlapping tumor cells. Moreover, manual counting was performed in 40 nonoverlapping tumor cells when: the average HER2 copy number per tile was ≥4.0 and ≤6.0 at the automated image analysis; automated counting of signals from the invasive cancer cells was impeded (low cellular density or high stromal density); polysomy or monosomy of chromosome 17 as defined by Tubbs and collaborators (20) was suspected. In addition, all cores with ratio >1.5 and <3.0 at the automated analysis were visually verified. Each informative core was evaluated separately in a blind fashion. Cores were analyzed by trained technologists, and all results were validated by breast pathologists. Average results of informative cores were considered. IHC and FISH results were reported according to the 2013 ASCO/CAP guidelines (5).

Statistical analysis. Only cases with at least one informative core for IHC or FISH were included in the analysis. The concordance between IHC and FISH was analyzed by comparing average IHC staining and average gene amplification results. Averages of HER2 status by IHC and FISH on diagnostic TMA sections were available for 498 out of the 554 total cases (89.9%). The average HER2 status by IHC on diagnostic TMA and random TMA were available for 106 cases out of the 116 selected cases (91.4%), whereas those by
FISH on diagnostic TMA and random TMA were available for 87.9% of cases (102/116). Average HER2 status by IHC and FISH on random TMA sections were available for 99 cases out of the 116 selected cases (85.3%). In order to evaluate the agreement between the two methods, positive, negative and overall concordance were calculated as previously published (18) following the recommendation of the 2013 ASCO/CAP guidelines (5). The level of agreement was also measured using the Cohen’s kappa test (agreement adjusted by chance). All analyses were performed using SAS software (version 9.1.3; SAS Institute, Inc., Cary, NC, USA).

Results

Concordance of HER2 status determined by IHC and FISH on diagnostic TMA section is summarized in Table I. The overall agreement between the two methods was 93.8% (kappa value=0.60). The positive agreement and the negative agreements were 69.2%, and 92.9%, respectively. Among the 423 cases considered as FISH non-amplified, protein overexpression was observed in one (0.2%) case. Five (7.7%) out of the 65 cases that showed gene amplification were negative at IHC.

We also assessed the concordance rates between HER2 status determined by IHC and FISH on TMA section according to the number of informative cores per case (one or two vs. three or four cores). The overall, positive and negative agreements for cases with one or two informative cores were lower (85.7%, 58.3% and 90.3%, respectively) than for cases with three or four evaluable cores (91.9%, 75.6% and 94.3%, respectively).

Table II shows the concordance rate of HER2 status determined by IHC and FISH on random TMA section. The overall agreement was 83.5% (kappa value=0.67). The positive and negative agreements were 71.4%, and 90.3%, respectively. Among the 62 cases considered FISH non-amplified, none was evaluated as 3+ for protein overexpression. Of the 35 cases that showed gene amplification, two (5.7%) were negative at IHC.

Concordance of HER2 status determined by FISH on diagnostic TMA section and random TMA section is summarized in Table III. In this analysis, we observed an overall agreement of 98.0%. The positive and negative agreements were 97.2%, and 98.5%, respectively. Of the 66 cases considered FISH non-amplified on diagnostic TMA section, one (1.5%) was evaluated as equivocal on random TMA section. One (2.8%) out of the 36 cases that showed gene amplification on diagnostic TMA section was equivocal for HER2 gene amplification on random TMA section.

Table IV presents the concordance of HER2 status determined by IHC on diagnostic TMA section and random TMA section. The overall agreement between the two methods was 93.6% (kappa value=0.86). The positive agreement and the negative agreements were 92.9%, and 93.9%, respectively. Among the 66 cases considered negative for protein overexpression on diagnostic TMA section, four (6.1%) were evaluated as equivocal on random TMA section. Two out of the 28 cases evaluated as positive for protein overexpression on diagnostic TMA section were considered equivocal on random TMA section.

Discussion

We observed a moderate concordance rate between the HER2 status determined by IHC and FISH on TMA section in 498 consecutive breast cancer specimens. In our hands, we recorded 93.8% concordance (kappa value of 0.60) between the two techniques when 2013 ASCO/CAP scoring criteria were used. Similar studies performed on breast cancer specimens using TMA section reported concordance rates ranging from 78.1% to 98.0% between the HER2 status obtained by IHC and FISH (21-27). The concordance results observed in those studies were evaluated using the 2007 ASCO/CAP scoring criteria (6). When we analyzed our data
using the 2007 ASCO/CAP scoring criteria, overall concordance was 91.5% (453/495) (sensitivity was 64.1% (of the 64 amplified cases, six were IHC 0/1+, 17 were IHC 2+, and 41 were IHC 3+), specificity was 95.6% (of the 431 non-amplified cases, 412 were IHC 0/1+, 18 were IHC 2+ and one was IHC 3+) and kappa value was 0.66), within the concordance range reported in the above mentioned studies. Among the 431 cases considered FISH non-amplified, one (0.2%) was evaluated as positive for protein expression. Six (9.4%) of the 64 cases that showed gene amplification were negative at IHC.

To our knowledge, this is the first study that analyzed the concordance between IHC and FISH on TMA section in breast cancer specimens according to the 2007 and 2013 ASCO/CAP criteria. To date, only one study analyzed the concordance between IHC and FISH on whole-tissue sections in 189 non-consecutive breast cancer specimens using both ASCO/CAP scoring criteria (28). Similar to Garbar and collaborators, we recorded a decrease in the percentage of cases evaluated as FISH-amplified and IHC-negative when the 2013 ASCO/CAP guidelines were used. However, Garbar et al. reported one case considered FISH-non-amplified and IHC-positive using the 2013 ASCO/CAP scoring criteria versus no cases using the 2007 ASCO/CAP scoring criteria, whereas in our study, the utilization of the recent ASCO/CAP scoring criteria did not change the percentage of these cases.

We recorded 7.7% of amplified cases among those considered negative at IHC (9.4% when 2007 ASCO/CAP scoring criteria were used). Similar studies comparing the concordance between IHC and FISH on TMA section reported between 1.24% and 11.5% of IHC negative cases being FISH-amplified (21-23, 25, 27). In our study, one non-amplified case was considered HER2-overexpressing with IHC, representing thus 0.23% of the IHC-positive cases (same proportion when 2007 ASCO/CAP guidelines were used). Similar studies have reported discordance rates for this category ranging from 0% to 23.5% (21-23, 25). We noted that this discordant case showed polysomy of chromosome 17. In agreement with previous studies (21, 26, 27), we postulate that the presence of chromosome 17 aneuploidy could at least in part explain the inconsistence between results.

A goal of our project was also to evaluate the impact of the new ASCO/CAP guidelines on the evaluation of HER2 status. The updated guidelines recommend performing an initial test (IHC or ISH) on core biopsy and specimens should be retested with an alternative assay when results are equivocal (5). Since IHC and FISH were not routinely performed on core biopsy for the 554 consecutive breast cancer specimens, for 116 selected breast cancer specimens we performed IHC and FISH on an additional paraffin block, randomly chosen from among all paraffin blocks of the same specimen that presented the same histological type as the routine diagnostic paraffin block, i.e. the paraffin block that was used to make the diagnosis. Since this additional block was randomly chosen, we consider this block a proxy of the core biopsy performed by the radiologist under ultrasound control (that is a random selection of a tumor area). We then compared IHC and FISH results obtained on the random TMA section (which simulate the HER2 status that would be obtained when IHC and FISH were performed on core biopsy) with IHC and FISH results obtained on diagnostic TMA section (which represent results obtained on excisional breast cancer specimens). We recorded an excellent overall concordance (98.0%, kappa value=0.94) between the HER2 gene amplification status observed on diagnostic TMA section and on the random TMA section. The overall concordance rate between HER2 protein overexpression obtained on diagnostic TMA section and random TMA section, however, was 93.6%, and therefore does not fulfill the ASCO/CAP suggestion of concordance greater than 95%
for clearly negative and positive cases (5). Overall concordance rates between IHC and FISH performed on random TMA section and on diagnostic TMA section were also both lower than 95% (83.5% vs. 93.8%).

In a similar study conducted on a cohort of 139 breast cancer cases for which HER2 status was determined in more than one block of a single tumor focus, Bethune et al. analyzed the HER2 concordance between the different blocks (both determined on whole tissue section) and reported 96.4% concordance (29). Results were considered concordant when the final HER2 status of all blocks was the same, regardless of whether HER2 status was determined by IHC or FISH. Selected cases of our cohort are comparable to those of the mentioned study, since blocks were from different tumor foci for only one patient of the subset. When HER2 status was defined according to the combined IHC and FISH results, we found 93.7% (89/95) concordance between diagnostic block and randomly selected block [sensitivity was 94.6%; of the 37 cases evaluated as positive on the diagnostic TMA section, two were considered equivocal and 35 positive on the random TMA section], specificity was 93.1% (of the 58 cases considered negative on the diagnostic TMA section, 54 were considered negative, and four equivocal on the random TMA section) and the kappa-value was 0.85. This difference might be explained by the fact that Bethune et al. used the former ASCO/CAP scoring guidelines. Similar to Bethune et al., we observed that the majority of discordant cases (seven out of nine) displayed an equivocal HER2 immunostaining. Unlike our study, Bethune et al. did not differentiate whether the concordance rate between the blocks varied according to which test had been used to determine HER2 status (IHC or FISH), as in their study HER2 status was determined by IHC, and FISH had mainly been performed on specimens with equivocal IHC staining. In our hands, HER2 concordance between the different blocks was higher when HER2 status was determined by FISH than by IHC. A study performed on a cohort of breast cancer specimens with equivocal immunostaining (evaluated on whole-tissue section according to the scoring guidelines of the HercepTest) reported that 68% of cases had a different score when the IHC staining was performed on additional blocks from the same breast tumor specimen or from axillary lymph node metastasis (30). This discordance in immunological staining between different blocks was substantially higher compared to that in our study. This difference may be explained by the fact that Lewis et al. analyzed the concordance for IHC scoring exclusively in cases that were originally evaluated as equivocal at IHC. The interpretation of immunostaining being based on semi-quantitative scoring, reported that interobserver variability and therefore discrepancies in HER2 IHC results were particularly elevated for cases scoring 2+ (31).

Analogous studies performed on multifocal and multicentric breast cancer specimens have analyzed concordance of HER2 status (determined on whole-tissue section) between different blocks from different tumor foci (32-34). Compared to our results, two studies had observed slightly lower concordance rate in HER2 gene amplification determined on different blocks, ranging from 90.3% to 94% (33, 34). These studies, however, reported their results using the 2007 ASCO/CAP guidelines (34) or other criteria for HER2 gene amplification (33). Another study reported 93.5% concordance rate in HER2 status determined by IHC and FISH from different blocks, where the largest focus displayed the most positive result in 98.4% of cases (32). This latter study, however, did not compare IHC and FISH results separately.

In contrast to our study, an analogous study that compared HER2 status determined on the needle core biopsy and subsequent excisional biopsy (whole tissue section) of the same tumor by IHC and FISH observed a higher concordance rate when HER2 status was evaluated by IHC (98% vs. 92%) (35).

TMA technology has several advantages over the traditional method, including batch variability reduction and decrease in reagent and technical time required for staining (16, 17, 36, 37). The amount of tissue needed is also reduced (37). In addition, visual assessment of immunohistochemical staining might be easier on TMA slides comparatively to whole tissue sections as the evaluator can compare staining intensities from different specimens on the same TMA slide (38).

Although TMAs represent a useful tool for rapid and efficient examination of large numbers of tumor tissues, this technique presents some limitations, including the training of highly qualified technicians and core losses (16, 38). Furthermore, it has been criticized that TMA sections may not accurately represent histopathological characteristics of the whole-tissue section. However, we and others have demonstrated that even one to two 0.6 mm cores per case can reliably reproduce results achieved on whole-tissue sections (18, 39).

Conclusion

In conclusion, whereas HER2 gene amplification seems to be very constant in different tumor regions, we observed greater variability regarding HER2 IHC staining. Indeed, we recorded a very good concordance (98%) between HER2 gene amplification status determined on diagnostic TMA section and on random TMA section. Concordance between HER2 protein overexpression observed on diagnostic TMA section and on random TMA section was 93.6%, lower than the 95% concordance suggested by the ASCO/CAP. The recent guidelines for HER2 testing in breast cancer developed by the ASCO/CAP recommend effectuating an initial test with IHC or ISH in core biopsy specimen. If test
results are equivocal, reflex testing on tumor specimen section with an alternative assay (IHC or ISH) should be performed. The guidelines do not recommend retesting when the assay result is clearly negative or positive. Nonetheless, based on our observation, we suggest that HER2 status determined on core biopsy by IHC should be interpreted carefully and a confirmatory repeat test using ISH methods on core biopsy or on a tumor specimen section should be considered. In our hands, the FISH method allowed for reliable evaluation of HER2 status on core biopsy.

Conflicts of Interest and Funding

This work was supported by Hoffmann-La Roche Limited. The funding source had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The Authors express special thanks to Drs. Anne Choquette, Michel Beauchemin, Mohamed Amin Hashem, Sophie Laberge, Mohib Morcos, Nathalie Mourad, Alexandre Odashiro, and Ion Popa. The Authors are grateful to the personnel of the Service de Pathologie, especially to Céline Plourde, for their precious technical support. DF received doctoral fellowships from the Fonds de recherche du Québec - Santé (FRQS) and the Laval University Cancer Research Center. CD is a recipient of the Canadian Breast Cancer Foundation-Canadian Cancer Society Capacity Development award (award #703003) and the FRQS Research Scholar. This study was supported by the Fondation des Hôpitaux Enfant-Jésus – St-Sacrement. Clinical specimens were provided by the Fondation du cancer du sein du Québec and the Banque de tissus et de données of the Réseau de recherche sur le cancer of the FRQS, which is affiliated with the Canadian Tumour Repository Network.

References


Received March 31, 2017
Revised April 16, 2017
Accepted April 19, 2017