# Immunohistochemistry and fluorescence in situ hybridization assessment of HER2 in clinical trials of adjuvant therapy for breast cancer (NCCTG N9831, BCIRG 006, and BCIRG 005)

Edith A. Perez · Michael F. Press · Amylou C. Dueck · Robert B. Jenkins ·
Chungyeul Kim · Beiyun Chen · Ivonne Villalobos · Soonmyung Paik ·
Marc Buyse · Anne E. Wiktor · Reid Meyer · Melanie Finnigan ·
JoAnne Zujewski · Mona Shing · Howard M. Stern · Wilma L. Lingle ·
Monica M. Reinholz · Dennis J. Slamon

**Abstract** A comprehensive, blinded, pathology evaluation of HER2 testing in HER2-positive/negative breast cancers was performed among three central laboratories. Immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) analyses were performed on 389 tumor blocks from three large adjuvant trials: N9831, BCIRG-006, and BCIRG-005. In 123 cases, multiple blocks were examined. HER2 status was defined according to FDA-approved guidelines and was independently re-assessed at each site. Discordant cases were adjudicated at an on-site, face-to-face meeting. Results across three independent pathologists were concordant by IHC in 351/381 (92 %) and FISH in 343/373 (92 %) blocks. Upon adjudication, consensus was reached on 16/30 and 18/30 of discordant IHC and FISH cases, respectively, resulting in overall concordance rates of 96 and 97 %. Among 155 HER2-negative blocks, HER2 status was confirmed in 153 (99 %). In the subset of 102 HER2-positive patients from N9831/BCIRG-006, primary blocks from discordant cases were selected, especially those with discordant test between local and central laboratories. HER2 status was confirmed in 73 (72 %) of these cases. Among 118 and 113 cases with IHC and FISH results and >1 block evaluable, block-to-block variability/heterogeneity in HER2 results was seen in 10 and 5 %, respectively. IHC−/FISH− was confirmed for 57/59 (97 %) primary blocks from N9831 (locally positive, but centrally negative); however, 5/22 (23 %) secondary blocks showed HER2 positivity. Among 53 N9831 patients with HER2-normal disease adjudicated as IHC−/FISH— (although locally positive), there was a non-statistically significant improvement in disease-free survival

E. A. Perez (✉)
Mayo Clinic, 4500 San Pablo Road S., Jacksonville,
FL 32224, USA
e-mail: perez.edith@mayo.edu

M. F. Press · I. Villalobos
Norris Comprehensive Cancer Center,
University of Southern California, 1441 Eastlake Avenue,
Los Angeles, CA 90033, USA

A. C. Dueck
Mayo Clinic, 13400 E. Shea Blvd., Scottsdale, AZ 85259, USA

R. B. Jenkins · B. Chen · A. E. Wiktor · R. Meyer ·
W. L. Lingle · M. M. Reinholz
Mayo Clinic, 200 First St. SW, Rochester, MN 55905, USA

C. Kim · S. Paik · M. Finnigan
National Surgical Adjuvant Breast and Bowel Project, East
Commons Professional Building, 4 Allegheny Center-Floor 5,
Pittsburgh, PA 15212, USA

M. Buyse
IDDI, Belgium, Avenue Provinciale, 30,
1340 Louvain-la-Nueve, Belgium

J. Zujewski
National Cancer Institute, 6116 Executive Blvd.,
Suite 300, Bethesda, MD 20892, USA

M. Shing · H. M. Stern
Genentech, Inc., 1 DNA Way, South San Francisco,
CA 94080, USA

D. J. Slamon
University of California, Los Angeles, 405 Hilgard Avenue,
Los Angeles, CA 90095, USA

with concurrent trastuzumab compared to chemotherapy alone (adjusted hazard ratio 0.34; 95 % CI, 0.11–1.05; $p = 0.06$). There were similar agreements for IHC and FISH among pathologists (92 % each). Agreement was improved at adjudication (96 %). HER2 tumor heterogeneity appears to partially explain discordant results in cases initially tested as positive and subsequently called negative.

**Keywords**  Breast cancer · HER2 testing · FISH · IHC · Concordance

## Introduction

The identification of the human epidermal growth factor receptor-2 (HER2) as an important cellular marker in the pathophysiology and treatment of breast cancer [1–3] has highlighted the importance of reliable testing methodology [1, 2]. Multiple discussions and publications related to this issue have been presented, not only addressing type of test, reliability, and definition of "positivity," but also which tests may best help predict the efficacy of anti-HER2 therapies for patients [3–15]. Both North Central Cancer Treatment Group (NCCTG) and National Surgical Adjuvant Breast and Bowel Project (NSABP) NCI-supported Cancer Cooperative Groups demonstrated that approximately 3–7 % of breast cancers formerly assessed as HER2-positive in local laboratories were called HER2-normal (IHC < 10 % of cells with circumferential membrane staining; FISH HER2:-CEP17 ratio < 2.0) when evaluated centrally [11, 12, 16]. Patients with these HER2-normal breast cancers appeared to have similarly prolonged disease-free survival (DFS) when treated with trastuzumab as those HER2-positive breast cancers corroborated in the central laboratories [11–13, 16]. These findings could be correct or alternatively due to a wide range of methodologic variables, including differing cutoffs for positivity, reading errors, discordance between pathologist interpretation and/or intratumoral heterogeneity. To address these critical aspects of HER2 testing, we conducted a round-robin study designed to evaluate HER2 testing and impact on patient outcomes using samples from two fully annotated specimen banks. The first set included samples from one of the two NCI-sponsored HER2-positive adjuvant trials, NCCTG-N9831 that first reported these findings (Perez et al. [12]). The second set was from two separate Breast Cancer International Research Group adjuvant trials in HER2-positive (BCIRG-006) and HER2-normal (BCIRG-005) breast cancers [17, 18]. There were four major objectives of this study. The first was to determine the concordance in testing and reading shared samples for HER2 results by the three central laboratories of the NCCTG, BCIRG, and NSABP. The second was to determine whether discordant cases could be adjudicated in a face-to-face

meeting of all the pathologists simultaneously reading discordant samples at a multi-headed microscope. The third was to evaluate the presence and frequency of intratumoral heterogeneity in HER2 status across multiple blocks from the same patient. Fourth, we sought to determine the impact of trastuzumab therapy in patients whose tumors initially tested HER2-positive in local laboratories but found to be HER2-normal by central testing and subsequently adjudicated as HER2-normal in this round-robin study of N9831 cases.

## Methods

### Specimens

Primary tumor blocks ($n = 389$) from three adjuvant trials (NCCTG N9831/NCT00005970; BCIRG-005/NCT00312 208; BCIRG-006/NCT00021255; clinicaltrials.gov) were sampled from the respective central tissue banks of the two groups (Mayo Clinic, Rochester, MN and University of Southern California, Los Angeles, CA). Blocks from NSABP B-31 were not made available for this project. Patient tissues were randomly selected from cases with available blocks in prospectively defined subgroups as previously determined by each group's independent central review. The subgroups included tumors read centrally as IHC−/FISH−, IHC−/FISH+, IHC+/FISH−, and IHC+/FISH+. Additional criteria were availability of multiple tumor blocks from the same case and treatment (N9831 specimens only). The blocks included 86 blocks from 62 patients in N9831 whose disease was centrally classified as HER2-normal by IHC (IHC−; IHC 0, 1+, or 2+) and FISH-negative (FISH−; HER2:CEP17 ratio < 2.0) (Table 1). Disease outcome for these 62 N9831 patients, previously classified as HER2-positive in local testing, was available for this study. We also included 105 blocks from 51 N9831 and 18 BCIRG-006 patients whose HER2 results were discordant between IHC and FISH testing within the central laboratories (33 IHC+/FISH−, 36 IHC−/FISH+). Additional cases included 54 blocks from 37 patients whose disease was centrally IHC+/FISH+ from N9831 and BCIRG-006; and 144 blocks from 96 patients whose disease was centrally IHC−/FISH—from BCIRG-005. For 121 patients, two blocks from the same primary tumor site were examined, and for two patients, three blocks were examined (Table 1). This study was approved by the Institutional Review Boards of each participating laboratory.

### Specimen de-identification

The NCCTG and BCIRG submitted their respective tissue blocks to a central laboratory (Mayo Clinic) where the blocks were sectioned and re-identified, as described in the

**Table 1** Patient/disease characteristics

| | IHC−/FISH−/BCIRG 005 (N = 96) | IHC−/FISH−/N9831 (N = 62) | IHC−/FISH+ (N = 36) | IHC+/FISH− (N = 33) | IHC+/FISH+ (N = 37) | Total (N = 264) |
|---|---|---|---|---|---|---|
| **Trial** | | | | | | |
| BCIRG 005 | 96 (100 %) | – | – | – | – | 96 (36.4 %) |
| BCIRG 006 | – | – | 18 (50 %) | – | 18 (48.6 %) | 36 (13.6 %) |
| N9831 | – | 62 (100 %) | 18 (50 %) | 33 (100 %) | 19 (51.4 %) | 132 (50 %) |
| **Number of FFPE tumor blocks** | | | | | | |
| 1 | 48 (50 %) | 38 (61.3 %) | 18 (50 %) | 17 (51.5 %) | 20 (54.1 %) | 141 (53.4 %) |
| 2 | 48 (50 %) | 24 (38.7 %) | 18 (50 %) | 14 (42.4 %) | 17 (45.9 %) | 121 (45.8 %) |
| 3 | 0 (0 %) | 0 (0 %) | 0 (0 %) | 2 (6.1 %) | 0 (0 %) | 2 (0.8 %) |
| **Treatment** | | | | | | |
| Treated on BCIRG 005 | 96 (100 %) | – | – | – | – | 96 (36.4 %) |
| AC → T | – | 29 (46.8 %) | 13 (36.1 %) | 20 (60.6 %) | 15 (40.5 %) | 77 (29.2 %) |
| AC → TH | – | 33 (53.2 %) | 21 (58.3 %) | 13 (39.4 %) | 14 (37.8 %) | 81 (30.7 %) |
| TCH | – | – | 2 (5.6 %) | – | 8 (21.6 %) | 10 (3.8 %) |
| **Age** | | | | | | |
| Median | 47.5 | 48.0 | 51.0 | 53.0 | 47.0 | 49.0 |
| Range | (25.0–70.0) | (27.0–69.0) | (29.0–72.0) | (29.0–76.0) | (25.0–73.0) | (25.0–76.0) |
| **Race** | | | | | | |
| Not collected (BCIRG studies) | 96 (100 %) | – | 18 (50 %) | – | 18 (48.6 %) | 132 (50 %) |
| Other | – | 11 (17.7 %) | 2 (5.6 %) | 3 (9.1 %) | 1 (2.7 %) | 17 (6.4 %) |
| White | – | 51 (82.3 %) | 16 (44.4 %) | 30 (90.9 %) | 18 (48.6 %) | 115 (43.6 %) |
| **Menopausal status** | | | | | | |
| Pre | 58 (60.4 %) | 35 (56.5 %) | 21 (58.3 %) | 15 (45.5 %) | 19 (51.4 %) | 148 (56.1 %) |
| Post | 38 (39.6 %) | 27 (43.5 %) | 15 (41.7 %) | 18 (54.5 %) | 18 (48.6 %) | 116 (43.9 %) |
| **ER/PgR status** | | | | | | |
| ER or PgR positive | 70 (72.9 %) | 50 (80.6 %) | 26 (72.2 %) | 21 (63.6 %) | 17 (45.9 %) | 184 (69.7 %) |
| Other | 26 (27.1 %) | 12 (19.4 %) | 10 (27.8 %) | 12 (36.4 %) | 20 (54.1 %) | 80 (30.3 %) |
| **Primary surgery** | | | | | | |
| Lumpectomy | 22 (22.9 %) | 26 (41.9 %) | 11 (30.6 %) | 12 (36.4 %) | 11 (29.7 %) | 82 (31.1 %) |
| Quadrantectomy/segmental | 25 (26 %) | 0 (0 %) | 4 (11.1 %) | 0 (0 %) | 4 (10.8 %) | 33 (12.5 %) |
| Mastectomy | 49 (51 %) | 36 (58.1 %) | 21 (58.3 %) | 21 (63.6 %) | 22 (59.5 %) | 149 (56.4 %) |
| **Nodal status** | | | | | | |
| Node positive (1–3 nodes) | 52 (54.2 %) | 31 (50 %) | 11 (30.6 %) | 14 (42.4 %) | 10 (27 %) | 118 (44.7 %) |
| Node positive (4–9 nodes) | 30 (31.3 %) | 16 (25.8 %) | 6 (16.7 %) | 10 (30.3 %) | 12 (32.4 %) | 74 (28 %) |
| Node positive (10+ nodes) | 14 (14.6 %) | 9 (14.5 %) | 4 (11.1 %) | 3 (9.1 %) | 6 (16.2 %) | 36 (13.6 %) |

**Table 1** continued

| | IHC−/FISH−/BCIRG 005 (N = 96) | IHC−/FISH−/N9831 (N = 62) | IHC−/FISH+ (N = 36) | IHC+/FISH− (N = 33) | IHC+/FISH+ (N = 37) | Total (N = 264) |
|---|---|---|---|---|---|---|
| Positive sentinel node | 0 (0 %) | 3 (4.8 %) | 2 (5.6 %) | 3 (9.1 %) | 2 (5.4 %) | 10 (3.8 %) |
| Node negative | 0 (0 %) | 3 (4.8 %) | 13 (36.1 %) | 3 (9.1 %) | 7 (18.9 %) | 26 (9.8 %) |
| Predominant tumor histology | | | | | | |
| Ductal | 80 (83.3 %) | 57 (91.9 %) | 29 (80.6 %) | 32 (97 %) | 32 (86.5 %) | 230 (87.1 %) |
| Lobular | 7 (7.3 %) | 3 (4.8 %) | 3 (8.3 %) | 1 (3 %) | 1 (2.7 %) | 15 (5.7 %) |
| Other | 9 (9.4 %) | 2 (3.2 %) | 4 (11.1 %) | 0 (0 %) | 4 (10.8 %) | 19 (7.2 %) |
| Histologic tumor grade (Elston/SBR) | | | | | | |
| Missing | 5 (5.2 %) | 3 (4.8 %) | 0 (0 %) | 2 (6.1 %) | 1 (2.7 %) | 11 (4.2 %) |
| Poor | 43 (44.8 %) | 35 (56.5 %) | 23 (63.9 %) | 20 (60.6 %) | 27 (73 %) | 148 (56.1 %) |
| Well/intermediate | 48 (50 %) | 24 (38.7 %) | 13 (36.1 %) | 11 (33.3 %) | 9 (24.3 %) | 105 (39.8 %) |
| Tumor size (cm) | | | | | | |
| ≤2 | 27 (28.1 %) | 14 (22.6 %) | 14 (38.9 %) | 8 (24.2 %) | 10 (27 %) | 73 (27.7 %) |
| >2 | 69 (71.9 %) | 48 (77.4 %) | 22 (61.1 %) | 25 (75.8 %) | 27 (73 %) | 191 (72.3 %) |

online supplement. This re-identification process insured that neither the contributing organization nor the individual subject could be identified without use of a coding key that was retained by the central statistical office for the purpose of blinding all reading pathologists as to the original classification of cases.

HER2 testing methods

The IHC HercepTest[TM] kit was used to determine HER2 protein expression according to manufacturer's instructions (Dako, Carpenteria, CA). The FISH PathVysion® HER2 DNA probe kit/HER2/CEP17 probe mixture (Abbott Molecular, Des Plaines, IL) was used to determine HER2 gene and chromosome 17 copy number in each of the 60 nuclei with slight modifications by each laboratory (Supplemental Table 1). HER2-positivity was defined according to FDA-approved guidelines used in the clinical trials (IHC+: uniform, intense circumferential membrane staining in >10 % invasive tumor cells; FISH+: HER2/CEP17 ratio ≥ 2.0). The BCIRG central laboratory also required that the average HER2 copy number be ≥ 4.0 copies per tumor cell nucleus to be scored as amplified [19]; the other central laboratories did not have this requirement. The HER2 status (IHC: 3+ vs. 0–2+; FISH: amplified vs. not) for each block was independently determined at each site.

Adjudication

The IHC and FISH cases that were discordant among the three pathologists were adjudicated by the three groups at a face-to-face meeting to determine (in a blinded analysis) the presence or absence of the HER2 alteration. This analysis was in addition to the re-analysis that had already been performed on these cases by each individual pathology group. In cases where consensus was not reached, slides were re-assayed (stained and scored) by FISH at the University of Southern California, and the stained slide was then sent to the other two central laboratories for scoring. After completion of all pathology and adjudication activities, data were unblinded and analyzed by the statistician (ACD).

Statistical analysis

A block must have had two or more pathology reads to be included in a given analysis. The adjudicated score and the majority score of the remaining cases in which a consensus was not reached (N = 14 for IHC, N = 12 for FISH) were used in all statistical analyses. The percent agreement (HER2-positive vs. HER2-normal) between the original central review result and round-robin final HER2 result of the primary block (defined as the block used in the original central review for clinical trial eligibility; one per patient)

was computed in the subsets of BCIRG-005 HER2-normal, N9831 HER2-normal, and N9831/BCIRG-006 HER2-positive specimens (all by central review). The agreement among blocks within patients in the same subsets of specimens was also estimated using percent agreement.

Disease-free survival was defined as time from randomization to first local, regional, or distant recurrence, contra-lateral breast cancer, another primary cancer (except squamous or basal cell skin cancer, carcinoma in situ of the cervix, or lobular carcinoma in situ of the breast), or death from any cause. DFS of N9831 patients with local pathologist's positive HER2 result but central and adjudicated (on all blocks) HER2-normal (IHC−/FISH−) breast cancer, was plotted by arm using Kaplan–Meier curves and compared via a Cox proportional hazards model stratified by hormone receptor and nodal status. Statistical significance was defined as a two-sided $p$ value $< 0.05$ throughout.

### Role of the funding source

The funding source had no role in study design, data, or writing. The corresponding author, Dr. Edith Perez, had full access to all data in the study and had final responsibility for the decision to submit for publication.

## Results

### HER2 concordance and adjudication

Independent reads were concordant across the three pathologists by IHC in 351/381 (92 %) cases and by FISH in 343/373

(92 %) cases (Fig. 1). Consensus was reached on 16/30 discordant IHC and 18/30 discordant FISH cases (Fig. 1). Thus, adjudication led to consensus in 367/381 (96 %) and 361/373 (97 %) of IHC and FISH cases, respectively.

Fourteen (4 %) IHC and 12 (3 %) FISH cases could not be adjudicated after face-to-face review and after re-testing of the FISH cases. Only one case was common between these two sets. Nine of 14 (64 %) non-adjudicated IHC cases had a two-thirds majority IHC 2+ score. Nine of 12 (75 %) non-adjudicated FISH cases had a two-thirds majority of non-amplified. The 12 non-adjudicated FISH cases had HER2:CEP17 FISH ratios spanning the 2.0 cut-off: ranging from 1.54 to 2.36 (average: 2.01; NCCTG), 1.13–2.22 (average: 1.72; NSABP), and 1.43–2.45 (average: 1.92; BCIRG).

Of 373 blocks with both an adjudicated IHC and FISH result, the overall concordance between IHC and FISH was 92 % (343/373). Among the IHC-negative blocks, concordance with FISH-negativity was 94 % (264/281), and among the IHC-positive blocks, concordance with FISH-positivity was 86 % (79/92). Among FISH-negative blocks, concordance with IHC-negativity was 95 % (264/277), and among the FISH-positive blocks, concordance with IHC-positivity was 82 % (79/96).

### Concordance between round-robin and original central review result

In the primary block of 96 BCIRG-005 patients with HER2-normal disease, IHC and FISH-negativity were confirmed in all 96 (100 %) cases (Table 2). In the primary block of 59 evaluable N9831 central HER2-normal cases,



Fig. 1 Overall concordance. The number of blocks showing concordance/discordance in IHC and FISH testing among three central laboratories. *Retest: 19 of the original 30 discordant FISH cases were not adjudicated at the face-to-face meeting. These 19 cases were re-assayed (stained and scored) by FISH at USC, and the stained slide was then sent to the other two central laboratories for scoring
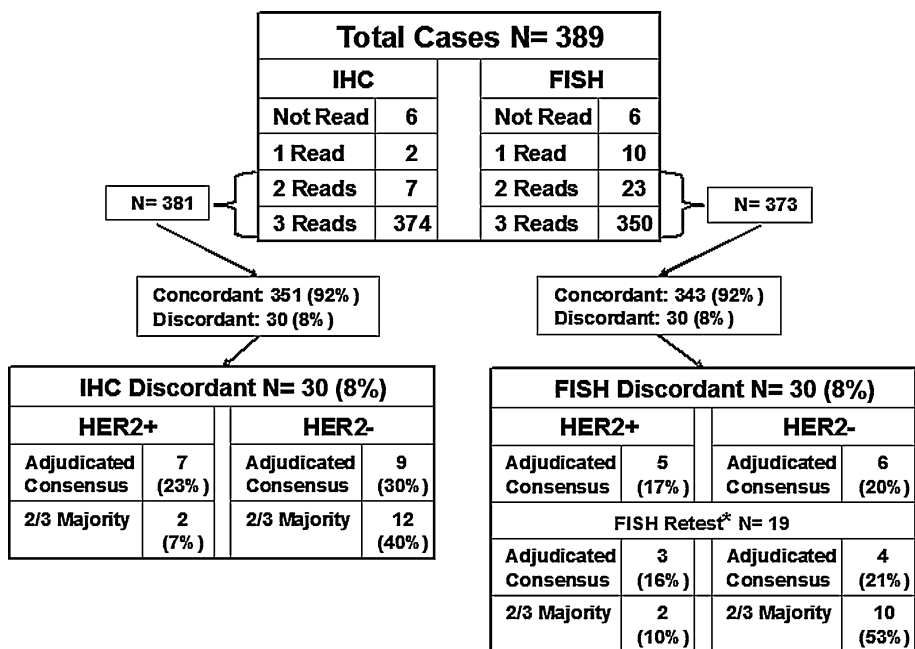
| Total Cases N= 389 | | | |
|---|---|---|---|
| **IHC** | | **FISH** | |
| Not Read | 6 | Not Read | 6 |
| 1 Read | 2 | 1 Read | 10 |
| 2 Reads | 7 | 2 Reads | 23 |
| 3 Reads | 374 | 3 Reads | 350 |

N= 381

N= 373

Concordant: 351 (92 %)
Discordant: 30 (8 %)

Concordant: 343 (92 %)
Discordant: 30 (8 %)

| IHC Discordant N= 30 (8%) | | | |
|---|---|---|---|
| **HER2+** | | **HER2−** | |
| Adjudicated Consensus | 7 (23%) | Adjudicated Consensus | 9 (30%) |
| 2/3 Majority | 2 (7%) | 2/3 Majority | 12 (40%) |

| FISH Discordant N= 30 (8%) | | | |
|---|---|---|---|
| **HER2+** | | **HER2−** | |
| Adjudicated Consensus | 5 (17%) | Adjudicated Consensus | 6 (20%) |
| FISH Retest* N= 19 | | | |
| Adjudicated Consensus | 3 (16%) | Adjudicated Consensus | 4 (21%) |
| 2/3 Majority | 2 (10%) | 2/3 Majority | 10 (53%) |

**Table 2** The concordance between central and adjudicated HER2 status in the primary block

| Study | Central | | | Adjudicated | | % |
|---|---|---|---|---|---|---|
| | HER2 status | N | | HER2 status | N | |
| BCIRG 005 | IHC−/FISH− | 96 | | IHC−/FISH− | 96 | 100 |
| N9831 | IHC−/FISH− | 59 | | IHC−/FISH− | 57 | 97 |
| | | | | IHC−/FISH+ | 2 | 3 |
| BCIRG 006 | IHC+/FISH+ | 36 | | IHC+/FISH+ | 34 | 94 |
| | | | | IHC−/FISH+ | 2 | 6 |
| | | | | IHC+/FISH− | 8 | 25 |
| N9831 | IHC+/FISH− | 32 | | IHC+/FISH+ | 4 | 13 |
| | | | | IHC−/FISH+ | 1 | 3 |
| | | | | IHC−/FISH− | 19 | 59 |
| | | | | IHC−/FISH+ | 6 | 18 |
| | IHC−/FISH+ | 34 | | IHC−/FISH− | 10 | 29 |
| | | | | IHC+/FISH+ | 17 | 50 |
| | | | | IHC+/FISH− | 1 | 3 |

Cases with only both IHC and FISH adjudicated results

IHC and FISH-negativity were confirmed in the primary block in 57 (97 %) cases (Table 2). In the primary block of 102 N9831/BCIRG-006 HER2-positive cases, HER2-positivity was confirmed in 73 (72 %) cases (Table 2), resulting from 29/102 primary blocks being either FISH-negative or IHC-negative locally, and FISH-negative and IHC-negative centrally. Among these 102 primary blocks, all 36 (100 %) BCIRG-006, originally HER2-positive blocks, were consistently evaluated as FISH-positive (the definition of "HER2-positive" in BCIRG-005/BCIRG-006 trials); and 34 (94 %) were also concordantly considered IHC-positive (IHC not assessed for entry to BCIRG trials) (Table 2). In the primary block of 66 cases with central discordant IHC/FISH status, the adjudicated result agreed in 14 (21 %) cases (Table 2).

Block-to-block intratumoral HER2 heterogeneity overall

Among 121 patients with two tissue blocks and two patients with three tissue blocks for analysis (Tables 1, 3), IHC results were obtained in 118 patients with the adjudicated IHC result agreeing across blocks in 106 (90 %) (Table 3). Among 113 patients with FISH results for >1 block, the adjudicated FISH result agreed across blocks in 107 (95 %) (Table 3). Of the 22 N9831 patients with HER2-normal (IHC-negative/FISH-negative) disease with duplicate blocks, 5/22 (23 %) tested positive (by IHC and/ or FISH) in at least one of the duplicate blocks (Table 3), clearly demonstrating heterogeneous HER2 gene copy number and/or protein immunostaining in the same tumor (Fig. 2). Moreover, the focal HER2-amplified region corresponded to the areas of HER2 protein over-expression, confirming the presence of a population of HER2-positive

cells in these tumors that had been called HER2-normal (Fig. 2a, b).

Block-to-block intratumoral HER2 heterogeneity and patient outcome in N9831 central and adjudicated HER2-normal cases

Of the original 103 patients identified in N9831 as locally HER2-positive but centrally HER2-normal, 59 blocks were evaluable out of 62 cases with tissue. Three cases were technical failures due to insufficient tumor tissue. Of the remaining 59 patients, 53 (90 %) had disease adjudicated as IHC-negative/FISH-negative for HER2. Among 22 of these 53 with more than one block available for analysis, 1/22 (5 %) was adjudicated as HER2-positive (by either IHC or FISH) in the primary block and 4/22 (18 %) had a second block adjudicated as HER2-positive (by either IHC or FISH) (Tables 2, 3). Among the 53 N9831 HER2-normal cases adjudicated as IHC-negative and FISH-negative (despite a prior HER2-positive test locally), there was an unadjusted trend in improvement in DFS associated with trastuzumab given concurrently with paclitaxel after doxorubicin/cyclophosphamide compared to chemotherapy alone (HR = 0.31, $p = 0.06$, 95 % CI 0.11–0.91; AC-T: 23 pts, ten events; AC-TH: 30 pts, five events). When adjusted for hormone receptor and nodal status, there was a non-statistically significant improvement in DFS associated with trastuzumab administered concurrently with chemotherapy compared to chemotherapy alone, based on a relatively small number of events (HR = 0.34, $p = 0.06$, 95 % CI 0.11–1.05; AC-T: 23 pts, ten events; AC-TH: 30 pts, five events) (Fig. 3). When considering only those few cases in which two blocks were both adjudicated as IHC-negative/FISH-negative ($n = 17/53$), a similar non-

**Table 3** Intratumoral heterogeneity

| Cohort | Adjudicated block 1** | Adjudicated block 2 | N | % |
|---|---|---|---|---|
| BCIRG 005 central HER2 | | | | |
| IHC−/FISH− (N = 47) | IHC−/FISH− | IHC−/FISH− | 47 | 100 |
| N9831 central HER2 | | | | |
| | IHC−/FISH− | IHC−/FISH− | 17 | 77 |
| | IHC−/FISH− | IHC+/FISH+ | 2 | 9 |
| IHC−/FISH− (N = 22) | IHC−/FISH− | IHC+/FISH− | 1 | 5 |
| | IHC−/FISH− | IHC−/FISH+ | 1 | 5 |
| | IHC−/FISH+ | IHC−/FISH− | 1 | 5 |
| BCIRG 006/N9831 central HER2 | | | | |
| IHC+/FISH+ (N = 17) | IHC+/FISH+ | IHC+/FISH+ | 16 | 94 |
| | IHC+/FISH+ | IHC−/FISH+ | 1 | 6 |
| | IHC+/FISH− | IHC+/FISH− | 3 | 21 |
| IHC+/FISH− (N = 14) | IHC+/FISH− | IHC−/FISH+ | 1 | 7 |
| | IHC+/FISH+ | IHC+/FISH+ | 1 | 7 |
| | IHC−/FISH− | IHC−/FISH− | 9 | 64 |
| | IHC+/FISH+ | IHC+/FISH+ | 5 | 39 |
| | IHC+/FISH+ | IHC−/FISH+ | 2 | 15 |
| IHC−/FISH+ (N = 13) | IHC+/FISH+ | IHC−/FISH− | 1 | 8 |
| | IHC+/FISH− | IHC−/FISH− | 1 | 8 |
| | IHC−/FISH− | IHC−/FISH− | 4 | 31 |

The IHC and FISH HER2 status of blocks from the same patient. Cases with only both IHC and FISH adjudicated results

**Adjudicated block 1 was the original centrally reviewed block used in the trial

significant trend in DFS improvement was observed (HR = 0.29, $p$ = 0.16, 95 % CI 0.05–1.65; AC-T: 8 pts, four events; AC-TH: 9 pts, two events).

## Discussion

The importance of HER2 as a prognostic marker in invasive breast cancer is well established [1, 2]. As such, it is critical to validate and standardize testing strategies to make an accurate assessment of HER2 status [3, 6]. The value of the current study becomes even more relevant with the provocative results from N9831 and NSABP-B31 demonstrating that in patients whose tumors were classified as HER2-normal based on central testing (although originally testing positive in local labs), there appears to be similar hazard ratios for benefit from adjuvant trastuzumab-based therapy when compared to chemotherapy alone [12, 16]. If patients exist who could benefit from this well-tolerated, effective therapeutic, but who may be misclassified by current HER2 testing (false-negatives), newer methods should be evaluated with attempts to determine better ways of identifying these patients. Similarly, we need to determine if there are patients receiving trastuzumab who are unlikely to benefit [9, 20]. The main objective of this study is to address critical aspects of HER2 testing through actual collaborative methodologic evaluation rather than a consensus review of literature published by independent groups [9]. To accomplish this, we conducted a round-robin study among pathologists from three central laboratories utilizing blocks from two HER2-positive adjuvant trastuzumab trials (N9831 and BCIRG-006) and one HER2-normal trial (BCIRG-005) to evaluate current HER2 testing methods and their potential impact on clinical outcomes in tumors from annotated trials.

The pre-round robin discordance rate for HER2 status (both IHC and FISH) in these cases as tested among the three expert pathologists was 8 %. At adjudication, a ≥96 % agreement was observed among these same pathologists, suggesting that interpretation issues and/or HER2 tumor heterogeneity may play a significant role in discordant results. The overall concordance between the adjudicated IHC and FISH results was 92 %.

Similar to the results of an international HER2 proficiency group study performed between five central laboratories [21], the majority of samples that could not be successfully adjudicated had IHC or FISH equivocal results as defined by the ASCO/CAP guidelines for HER2 positivity [9]. Of the 14 IHC cases that could not be successfully adjudicated, 64 % were classified as 2+. Circumferential distribution and character (intensity, granularity) of cell staining, rather than quantity of stained cells, was the main reason for discordance among the round-robin pathologists, similar to what was reported in a
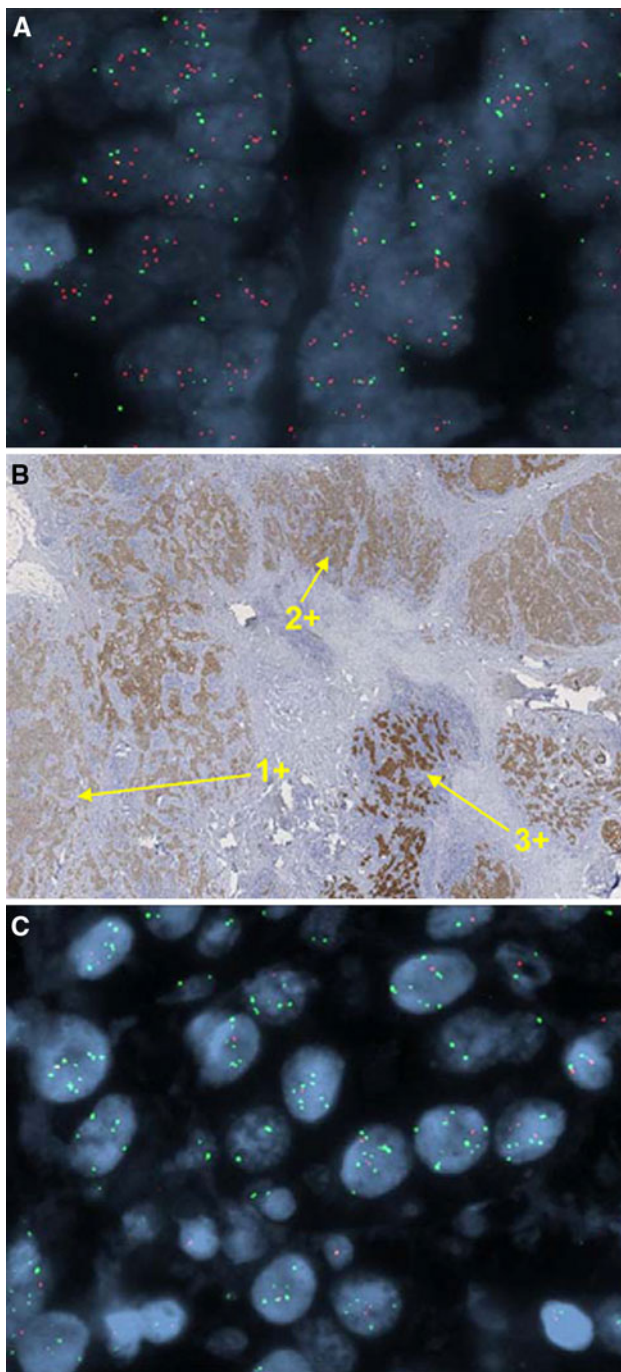
Fig. 2 Intratumoral HER2 heterogeneity. HER2 protein and gene/chromosome heterogeneity in the same tumor. **a** HER2 gene amplification. Representative FISH staining demonstrating a focal HER2 amplified clone that corresponds to the area of HER2 protein over-expression in **b**. **b** Variable HER2 IHC Protein Immunostaining. The area identified shows IHC 3+ immunostaining, while the remainder of the microscopic field shows IHC 2+ immunostaining heterogeneity. **c** Representative FISH staining demonstrating polysomy 17 in the same tumor as in **a** and **b**
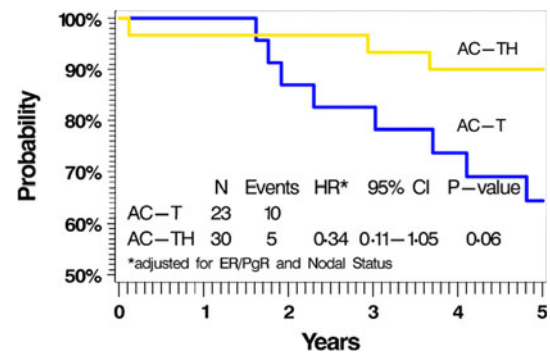


Fig. 3 Kaplan–Meier curves of DFS in N9831 patients with IHC−/FISH− disease. All patients had IHC−/FISH− disease by central review and all blocks adjudicated in the current study as IHC−/FISH−

copy number of 4.67 per nucleus across pathologists, both of which are near the FDA-approved cut-offs of 2.0 for ratio and 4.0 for copy number.

Although the absolute HER2 counts were similar across the reading pathologists, small changes in the CEP17 counts (denominator) can and did significantly affect this ratio and changed amplification status of HER2 when using the HER2:CEP17 ratio. As a result, the interpretation of HER2 gene amplification differed in these cases. These data indicate that in equivocal cases, HER2 gene and CEP17 copy numbers should be assessed independently and may have important clinical implications [23].

When the HER2 results are in the equivocal range, pathologists should consider consulting with a second pathologist to corroborate or possibly adjudicate the HER2 status. An accompanying explanation/interpretation of the HER2 status from the pathologist(s) is critical to help guide the clinician in making appropriate management decisions [4]. The patient also needs to be informed of challenges associated with HER2 testing, particularly in cases near the FISH ratio of 2.0 as recommended by the FDA. A trend toward benefit from trastuzumab (adjusted HR = 0.34; $p$-0.06) was observed in the small subset of N9831 patients ($n = 53$) with disease deemed HER2-normal by central review and confirmed on the limited number of blocks in the round-robin (although initially called HER2-positive locally). While this observation is based on a very small number of events, we recognize that trends toward benefit are important to document, despite failing to reach statistical significance. An alternative and equally plausible explanation for the observation of benefit in HER2-normal cases could be the heterogeneity we observed in both HER2 protein overexpression and gene amplification during assessment of more than one block from the same patient. The current data indicate that 5/22 patients (23 %) that had locally tested as HER2-positive in N9831 and subsequently called IHC-

recent HER2 proficiency testing study [22]. The 12 non-adjudicated FISH cases had HER2:CEP17 FISH ratios with an average of 1.88 (range 1.13–2.45) and an average HER2

negative/FISH-negative centrally, had a second block from the same primary tumor that we found to be HER2-positive in the round-robin analysis. Accordingly, testing an additional portion of the tumor from "HER2-normal" cases may be advisable to avoid the possibility of not treating patients who might benefit from HER2-targeted therapies. This phenomenon could account for some discrepancies between local and central HER2-testing in N9831 and NSABP B-31. One option to avoid this would be to consider testing a second section from an additional tumor block for each patient whose initial HER2 testing yields normal results.

The incidence of variable IHC staining was higher for overexpression (10 %) compared to heterogeneity of amplification status (5 %). Variable rates of intratumoral HER2 heterogeneity (<1–30 %) of unknown clinical significance have also been reported [15, 24–28]. In addition, we occasionally observed distinct intratumoral heterogeneity within the same tumor section (Fig. 2). When distinct populations of cells exist in the same section, the HER2 status (by IHC and FISH) should be reported in accordance with established guidelines [9].

Approximately 8–10 % of HER2-amplified breast cancers are falsely negative (IHC 0 or 1+) by IHC HER2 testing [14, 15], and given the high incidence of breast cancer in the US, this rate of false-negativity could negatively and critically impact the care of more than 5,000 patients each year [8]. Based on the approximate 50 % reduction of relapse events when anti-HER2-targeted therapy is included in the adjuvant setting, thousands of women in the US each year may be experiencing relapses from failed adjuvant regimens that may not have occurred had these women been correctly identified as patients with HER2-positive disease and initially treated with trastuzumab- or lapatinib-based regimens. Improving the accuracy of HER2 testing and reducing the incidence of false-negative results can directly save lives [29].

## References

1. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235(4785):177–182
2. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A et al (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science 244(4905):707–712
3. Sauter G, Lee J, Bartlett JM, Slamon DJ, Press MF (2009) Guidelines for human epidermal growth factor receptor 2 testing: biologic and methodologic considerations. J Clin Oncol 27(8): 1323–1333
4. Ahmed SS, Iqbal J, Thike AA, Lim AS, Lim TH, Tien SL, Tan PH (2011) HER2/neu revisited: quality and interpretive issues. J Clin Pathol 64(2):120–124
5. De P, Smith BR, Leyland-Jones B (2010) Human epidermal growth factor receptor 2 testing: where are we? J Clin Oncol 28(28):4289–4292
6. Shah S, Chen B (2010) Testing for HER2 in breast cancer: a continuing evolution. Pathol Res Int 2011:903202
7. Vogel UF (2010) Confirmation of a low HER2 positivity rate of breast carcinomas—limitations of immunohistochemistry and in situ hybridization. Diagn Pathol 5:50
8. Ross JR (2009) Breast cancer biomarkers and HER2 testing after 10 years of anti-HER2 therapy. Drug News Perspect 22(2): 93–106
9. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A et al (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25(1):118–145
10. Shah SS, Ketterling RP, Goetz MP, Ingle JN, Reynolds CA, Perez EA, Chen B (2010) Impact of American Society of Clinical Oncology/College of American Pathologists guideline recommendations on HER2 interpretation in breast cancer. Hum Pathol 41(1):103–106
11. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, Flynn PJ, Ingle JN, Visscher D, Jenkins RB (2006) HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol 24(19):3032–3038
12. Perez EA, Reinholz MM, Hillman DW, Tenner KS, Schroeder MJ, Davidson NE, Martino S, Sledge GW, Harris LN, Gralow JR et al (2010) HER2 and chromosome 17 effect on patient outcome in the N9831 adjuvant trastuzumab trial. J Clin Oncol 28(28): 4307–4315
13. Press MF, Slamon DJ, Flom KJ, Park J, Zhou JY, Bernstein L (2002) Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. J Clin Oncol 20(14):3095–3105
14. Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou JY, Wardeh R, Li YT, Guzman R, Ma Y et al (2005) Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. Clin Cancer Res 11(18):6598–6607
15. Press MF, Finn RS, Cameron D, Di Leo A, Geyer CE, Villalobos IE, Santiago A, Guzman R, Gasparyan A, Ma Y et al (2008) HER-2 gene amplification, HER-2 and epidermal growth factor receptor mRNA and protein expression, and lapatinib efficacy in

women with metastatic breast cancer. Clin Cancer Res 14(23): 7861–7870

16. Paik S, Kim C, Wolmark N (2008) HER2 status and benefit from adjuvant trastuzumab in breast cancer. N Engl J Med 358(13): 1409–1411

17. Eiermann W, Pienkowski T, Crown J, Sadeghi S, Martin M, Chan A, Saleh M, Sehdev S, Provencher L, Semiglazov V et al (2011) Phase III study of doxorubicin/cyclophosphamide with concomitant versus sequential docetaxel as adjuvant treatment in patients with human epidermal growth factor receptor 2-normal, node-positive breast cancer: BCIRG-005 trial. J Clin Oncol 29(29): 3877–3884

18. Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M, Mackey J, Glaspy J, Chan A, Pawlicki M et al (2011) Adjuvant trastuzumab in HER2-positive breast cancer. N Engl J Med 365(14):1273–1283

19. Press MF, Bernstein L, Thomas PA, Meisner LF, Zhou JY, Ma Y, Hung G, Robinson RA, Harris C, El-Naggar A et al (1997) HER-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas. J Clin Oncol 15(8):2894–2904

20. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N, Leiberman G, Slamon DJ (2005) Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. Clin Breast Cancer 6(3):240–246

21. Dowsett M, Hanna WM, Kockx M, Penault-Llorca F, Ruschoff J, Gutjahr T, Habben K, van de Vijver MJ (2007) Standardization of HER2 testing: results of an international proficiency-testing ring study. Mod Pathol 20(5):584–591

22. Grimm EE, Schmidt RA, Swanson PE, Dintzis SM, Allison KH (2010) Achieving 95% cross-methodological concordance in

HER2 testing: causes and implications of discordant cases. Am J Clin Pathol 134(2):284–292

23. Dal Lago L, Durbecq V, Desmedt C, Salgado R, Verjat T, Lespagnard L, Ma Y, Veys I, Di Leo A, Sotiriou C (2006) Correction for chromosome-17 is critical for the determination of true Her-2/neu gene amplification status in breast cancer. Mol Cancer Ther 5(10):2572–2579

24. Vance GH, Barry TS, Bloom KJ, Fitzgibbons PL, Hicks DG, Jenkins RB, Persons DL, Tubbs RR, Hammond ME (2009) Genetic heterogeneity in HER2 testing in breast cancer: panel summary and guidelines. Arch Pathol Lab Med 133(4):611–612

25. Brunelli M, Manfrin E, Martignoni G, Miller K, Remo A, Reghellin D, Bersani S, Gobbo S, Eccher A, Chilosi M et al (2009) Genotypic intratumoral heterogeneity in breast carcinoma with HER2/neu amplification: evaluation according to ASCO/CAP criteria. Am J Clin Pathol 131(5):678–682

26. Burrell RA, Juul N, Johnston SR, Reis-Filho JS, Szallasi Z, Swanton C (2010) Targeting chromosomal instability and tumour heterogeneity in HER2-positive breast cancer. J Cell Biochem 111(4):782–790

27. Hanna W, Nofech-Mozes S, Kahn HJ (2007) Intratumoral heterogeneity of HER2/neu in breast cancer—a rare event. Breast J 13(2):122–129

28. Lewis JT, Ketterling RP, Halling KC, Reynolds C, Jenkins RB, Visscher DW (2005) Analysis of intratumoral heterogeneity and amplification status in breast carcinomas with equivocal (2 +) HER-2 immunostaining. Am J Clin Pathol 124(2):273–281

29. Ross JS (2010) Human epidermal growth factor receptor 2 testing in 2010: does chromosome 17 centromere copy number make any difference? J Clin Oncol 28(28):4293–4295