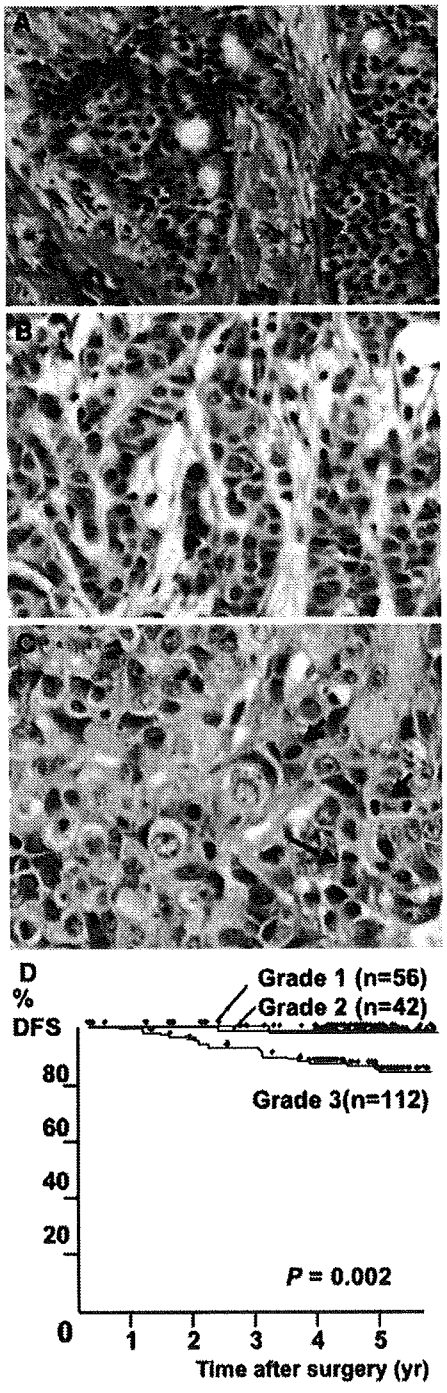

Fig. 5 Nuclear grade of pN0 (regional lymph nodes) invasive ductal carcinoma. Histopathological view of nuclear grade 1 (**a**), grade 2 (**b**), and grade 3 (**c**). **a** Nuclei are uniform, and mitotic figures are not seen. **c** Nuclei are pleomorphic, and there are a lot of mitotic figures (*arrows*). **b** is intermediate between **a** and **c**. H&E stain. ×200. **d** Disease-free survival curves for patients with pN0 breast cancer stratified by nuclear grades. Curves differ significantly ($P = 0.002$). *%DFS* Disease-free survival (%)

(Herceptin) therapy, and (4) high response rate to anthracycline-based chemotherapies [17].

Criteria for *HER2* overexpression and amplification have been established. *HER2* expression status, tested by IHC, is classified as score 0, 1+, 2+, or 3+ (Fig. 1 in [17]). An IHC score of 3+ is assessed to indicate overexpression, or *HER2*-positive; a score of 0 or 1+, as *HER2*-negative; a score of 2+, as equivocal, with a recommendation for retesting by fluorescence in situ hybridization (FISH).

There are two types of FISH tests, single-color FISH and dual-color FISH, but only the latter is approved for diagnostic testing in Japan. Dual-color FISH (PathVysion, Vysis/Abbott) visualizes concurrently red signals of *HER2* on 17q21.1 and green signals of *CEP17* on the centromere of chromosome 17. The sum of signals on 20 nuclei of cancer cells is counted for both *HER2* and *CEP17*, and the *HER2/CEP17* ratio is calculated by dividing the total signal number of *HER2* by that of *CEP17*. If the *HER2/CEP17* ratio is 2.0 or higher, the tumor is assessed to be FISH positive, whereas a ratio of less than 2.0 is taken to indicate that the tumor is FISH negative (Fig. 2 in [17]). *HER2*-positive cases are considered to be eligible for trastuzumab therapy [19].

In 2006, revised criteria for *HER2* gene amplification were recommended by the NCCN HER2 Testing in Breast Cancer Task Force as follows: a tumor with an IHC score of 0 or 1+, or with a *HER2/CEP17* ratio of less than 1.8 by dual-color FISH, is *HER2*-negative; a tumor with an IHC score of 3+, or with a *HER2/CEP17* ratio of more than 2.2 by dual-color FISH, is *HER2*-positive; a tumor with an IHC score of 2+ should be further tested using FISH, with its *HER2* status determined on the basis of the FISH result. Tumor samples with a *HER2/CEP17* ratio of 1.8 to 2.2 are considered to be borderline [20]. In Japan, a cut-off value of 2.0 for the *HER2/CEP17* ratio is still used.

Amplification of the *HER2* gene and overexpression of its protein is correlated with a poorer prognosis of patients with pN0 invasive breast cancer as well as patients with node-positive breast cancer. In a review, Ross et al. stated that *HER2* amplification or overexpression was correlated with poorer prognosis by univariate and/or multivariate analyses in 73 of 81 studies (25,166 of 27,161 patients) published between 1987 and 2003 [21]. Although *HER2* is frequently positive in DCIS of higher grade and with an accumulation of molecular alterations, i.e., comedo-type DCIS and Paget's disease, *HER2* is not a prognostic factor in DCIS.

A correlation between *HER2* overexpression and response to adjuvant or neoadjuvant anthracycline-based chemotherapy has been reported in many studies [22]. In the JBCRG-01 protocol, *HER2* overexpression was

confirmed to be a significant predictive factor of tumor response to anthracycline-based PST [23]. Hayes et al. [24] reported that a node-positive breast cancer showing amplification or overexpression of *HER2* can benefit from the addition of paclitaxel therapy after adjuvant treatment with doxorubicin plus cyclophosphamide, regardless of the ER status. In contrast, they considered that patients with *HER2*-negative, ER-positive, node-positive breast cancer would gain little benefit from the administration of paclitaxel after adjuvant chemotherapy with doxorubicin plus cyclophosphamide.

Lymphovascular involvement

Although lymphatic and vascular invasion (abbreviated as ly and v) is generally considered to be a parameter of aggressiveness of cancers in various organs, the significance of ly or v as a prognostic factor has been the subject of controversy because (1) ly(+) is frequently difficult to differentiate histologically from tumor nests in pseudolymphatic spaces emerging as artifacts, (2) ly(+) or v(+) can be confused with an intraductal component, and (3) pathologists tend to evaluate a cancer as ly(+) unconditionally if lymph node metastasis is present.

Lymphatic and vascular invasion have been excluded from the items routinely described in the "general rules". Nonetheless, obvious lymphatic permeation does exist and is sometimes diffuse in highly aggressive breast cancers, such as inflammatory carcinoma. Yoshimoto et al. [25] defined lymphovascular invasion as being strongly positive if four or more ly or v foci were present, positive if three or fewer ly or v foci were present, and negative if no ly or v focus was evident in all case slides. In their series of cases, the percentages of those that were ly or v strongly positive, positive, and negative were 13, 13, and 74%, respectively, and the corresponding recurrence rates were 60, 40, and 26%, respectively [25].

In the 2007 St Gallen meeting consensus, extensive peritumoral vascular invasion was considered to be one of the parameters for assessing whether pN0 breast cancer is of low risk or intermediate risk. In pN1a breast cancers, however, extensive lymphovascular invasion has been removed from the risk factors [1]. Extensive peritumoral vascular invasion is defined as the presence of neoplastic emboli in two or more blocks, which could be interpreted as both extensive ly(+) and v(+). Intratumoral lymph vessels in invasive breast cancer are suggested to be absent [26], and it is recommended that peritumoral lymphovascular invasion should be taken into account. Recently, several reports have mentioned the utility of D2-40 or vascular endothelial growth factor (VEGF)-C in combination with IHC to visualize lymphatic ducts, and the diagnostic application of these molecules is expected [27, 28].

Future perspectives

According to the recommendation of the St Gallen meetings, more than 70% of pN0 breast cancers are included in the intermediate-risk group and considered eligible for adjuvant chemotherapy, despite the fact that the 10-year disease-free survival rate of patients with pN0 breast cancer in Japan is 10–15% [4]. It is expected that recurrence risk in pN0 breast cancer will be evaluable more accurately based on the properties of the resected tumors. On the basis of estimations from the categories established at the St Gallen consensus meeting, the 10-year recurrence rate of pN0 intermediate-risk cancer would be 13–20%, whereas that of pN0 low-risk cancer would be 5% or less. It is ideal that the pN0 group is classified as representing 33% of the high-risk group, with a recurrence rate of 30–45%, and 67% of the low-risk group with a recurrence rate of 5% or less (Fig 6). To establish a more ideal classification of early breast cancer, various tests are being undertaken that are based on molecular features.

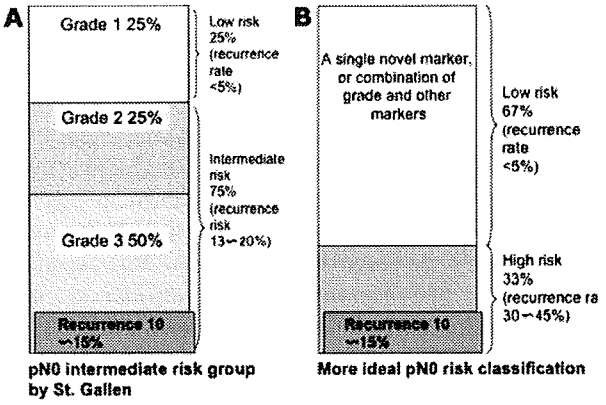

Fig. 6 Comparison between the pN0 intermediate-risk group (**a**) and the ideal pN0 high-risk group (**b**). In **a**, the intermediate-risk group comprises 75% of all patients with pN0 breast cancer and is estimated to show 13–20% of recurrence. Actually, most of pN0 breast cancers have an invasive component that is 2.0 cm or less in size. Lymphovascular invasion and *HER2* overexpression are shown by some researchers to be significant prognostic indicators with a relatively weak impact in pN0. Therefore, the overwhelming risk factor of pN0 breast cancer appears to be histological/nuclear grade. In **b**, the putative higher risk group comprises 33% of all patients with pN0 breast cancer and is estimated to show 30–45% of recurrence. For the concentration of cases, a single novel molecular marker or a combination of grade and other molecular markers would be effective. These molecular markers will be assayed by DNA microarray, Oncotype DX, array CGH, or other tools in the near future

DNA microarray

Using high-throughput microarray analysis, a 70-gene signature has been identified that can accurately select early-stage breast cancer patients who are highly likely to develop distant metastases and who, therefore, may obtain maximal benefit from adjuvant chemotherapy [29]. Validation and feasibility studies of the 70-gene profile (Mammaprint; Agendia) to patients of 60 years or younger with pN0 Stage I, II breast cancer are ongoing in the large adjuvant MINDACT (microarray in node-negative disease may avoid chemotherapy) clinical trial [29].

Oncotype DX

Although patients with pN0 and ER-positive breast cancer have an excellent prognosis, about 15% show relapse after 5 years of endocrine therapy. Clinical trials have provided evidence that these patients gain a significant benefit from chemotherapy; however, it would be significant overtreatment if it were applied to all of them. Strategies for evaluating tumors in a clinical setting have been developed using smaller sets of genes. One such strategy is the 21-gene assay (Oncotype DX; Genomic Health, Redwood City, CA), which is currently in commercial use in the USA. In Oncotype DX, a 21-gene recurrence score (RS) has been developed based on the monitoring of mRNA expression levels of 16 cancer-related genes in relation to five reference genes. One advantage of this test is the use of paraffin-embedded blocks in contrast to previous methods that required fresh frozen tissue. Oncotype DX has been shown to predict 10-year distant recurrence in patients with ER-positive, axillary lymph node-negative breast cancer. This genomic assay has also been shown to predict response to chemotherapy and endocrine therapy. A prospective study—the Trial Assigning Individualized Options for Treatment (Rx) (TAILORx)—to examine whether chemotherapy is required for the intermediate-risk group defined by the RS is accruing in North America [30, 31].

Array CGH (cancer array 800)

Comparative genomic hybridization (CGH) has already made a significant impact on cancer cytogenetics. In array-based CGH, DNAs spotted in a CGH-array contain sequence information directly linked to the genome database, and particular biological aspects of genes that lie within regions involved in copy-number aberrations can be noted. The application of array-based CGH technology for the diagnosis of breast cancer is awaited [32].

Basal-like type

With the use of gene microarrays, different subtypes of breast cancer have been characterized. These subtypes include the basal, ERBB2+, and luminal A, B and C subtypes [33, 34]. Basal-like-type breast cancer was determined by DNA microarray to be a group of tumors that were ER-negative, *HER2*-negative, and positive for myoepithelial/basal markers, such as vimentin, alpha-smooth muscle actin, EGFR, cytokeratin (CK) 5/6, CK14, and/or CK17 [33, 34]. It has also been reported that breast cancers arising in patients with familial breast/ovarian cancer carry germ-line *BRCA1* mutations [35].

There are at least four histologically distinct breast cancer groups with undifferentiated features, and these frequently show bidirectional differentiation toward luminal epithelial and myoepithelial/basal lineages [36, 37]. These groups comprise (1) invasive ductal carcinoma with a large central acellular zone (central acellular carcinoma), (2) atypical medullary carcinoma, (3) matrix-producing carcinoma, and (4) carcinoma with spindle-cell metaplasia (Fig. 7). In these four cancer types, KIT (CD117) expression and EGFR overexpression were detected in 34% and 88% of cancers with frequent expression of myoepithelial/basal markers but a low frequency of *HER2* overexpression or ER/PgR expression (Fig. 8) [37]. For the identification of a basal-like phenotype, confirmation of positivity for basal cytokeratins, i.e., either CK5/6 or CK14, by IHC is recommended [38].

Empirically, cases of node-negative breast cancer showing early recurrence appear to frequently contain the basal-like type [39]. However, Fulford et al. have reported that in node-negative patients, prognosis was similar between basal-like and other Grade 3 invasive ductal carcinomas, whereas basal-like grade 3 invasive ductal carcinoma showed a poorer prognosis than other types in patients who were positive for node metastasis [40].

One important current issue is the so-called triple-negative (i.e., ER/PgR-negative, *HER2*-negative) breast cancer, for which endocrine therapy or trastuzumab is not applicable. Approximately 5–15% of breast cancers are in this category. If systemic chemotherapies are not effective for "triple-negative" breast cancer, there are few treatment choices for this group and, in fact, chemotherapy is frequently not effective. A substantial percentage of "triple-negative" breast cancers appear to be of the basal-like type. The characterization of basal-like breast cancer on the basis of histological characteristics and molecular alterations would be useful for prognostication and treatment selection and also for the identification of targets for molecular therapy. It might be worth investigating whether therapies against activated KIT and/or EGFR are effective for cancers of the above-mentioned four histological types.

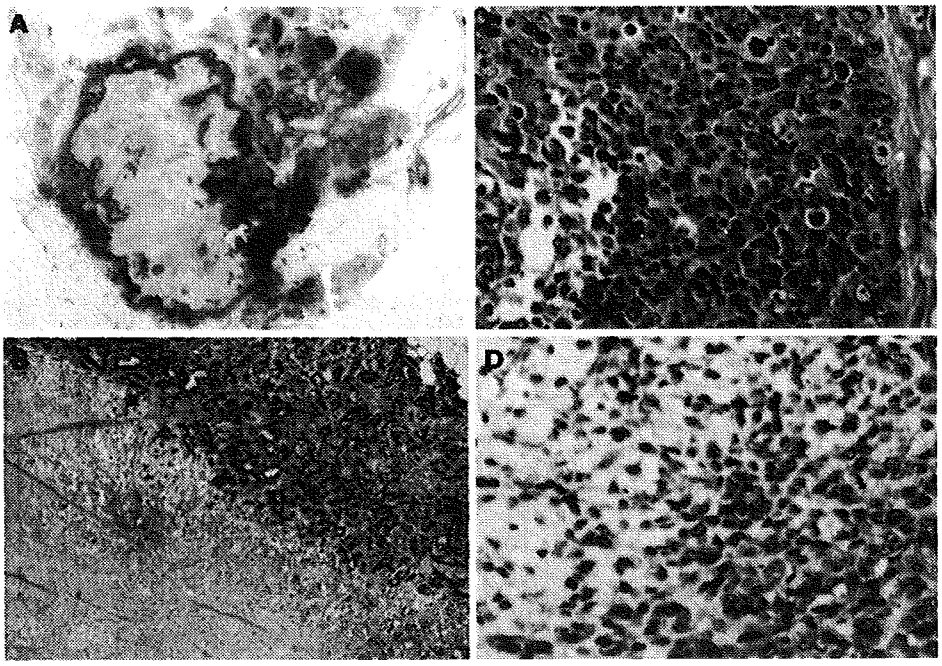

Fig. 7 Distinct breast cancer types of undifferentiated features, that appear to be representative of the basal-like phenotype. **a**, **b** Invasive ductal carcinoma with a large central acellular zone (central acellular carcinoma). **c** atypical medullary carcinoma, **d** matrix-producing carcinoma. Carcinoma with spindle-cell metaplasia is also included in this group. H&E stain. **a** ×1, **b** ×40, **c**, **d** ×200

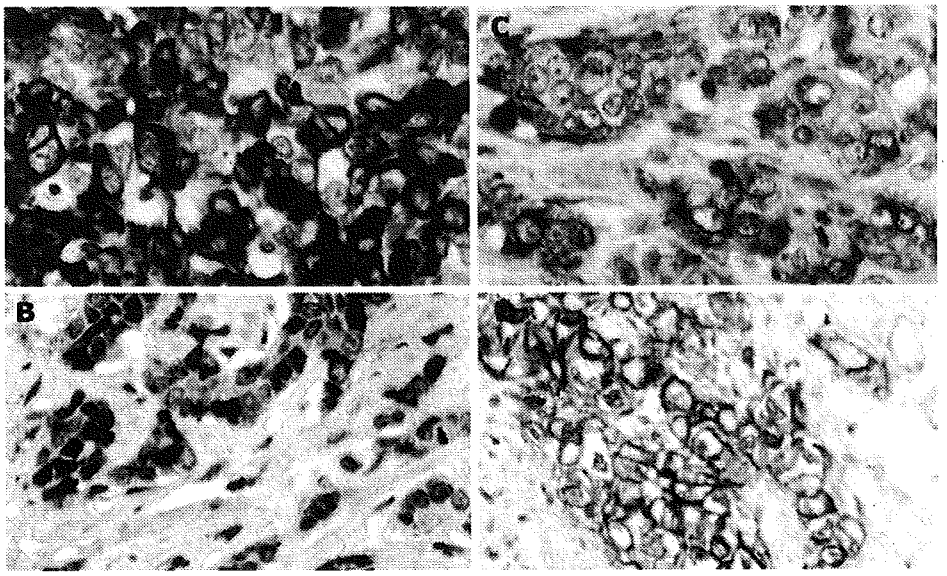

Fig. 8 Expression of myoepithelial/basal markers in undifferentiated-type breast carcinomas. **a** Vimentin, **b** KIT, **c** α-smooth muscle actin (SMA), **d** epidermal growth factor receptor (EGFR). The positive rates of expression of vimentin, α-SMA, KIT (CD117), and EGFR were 88, 41, 34, 88%, respectively, in undifferentiated types, whereas these were 5, 0.7, 5, and 5%, respectively, in other types [37]. Immunoperoxidase stain. ×200

Today, systemic chemotherapy, endocrine therapy, and trastuzumab are very effective, not only for metastatic breast cancer but also operable early breast cancer. Certain chemotherapeutic regimens with or without trastuzumab can achieve a pCR in the primary tumor in a proportion of cases. Nevertheless, these therapies are not perfect; they are not always effective, they may result in adverse events or late complications, and they may induce tumor resistance to the therapy in due course.

To achieve a greater decrease in the incidence of recurrence and death in patients with early breast cancer and those suffering adverse effects from the therapies, it is important to discriminate the higher risk group from the lower risk group more accurately in intermediate-risk

node-negative breast cancer. Furthermore, regardless of nodal status, the establishment of a diagnosis and treatment strategy for "triple-negative" breast cancers, especially the basal-like type, is desirable. To this end, molecular markers and tools, such as Oncotype Dx, Mammaprint, array CGH, and/or IHC markers for the basal-like type, would be effective. In addition, from a pathologist's viewpoint, proposals for the accurate evaluation of pT, pN, grade, ly, v, ER, PgR, and *HER2* have only recently been put forward. Therefore, there are few adequate follow-up data based on accurate descriptions of these pathological parameters. In the future, prospective clinical data as well as the revision of archival cases based on accurate histopathological evaluation might prove to have considerably higher value than expected hitherto, as exemplified by the proverb "it is always dark at the foot of the lighthouse".

Acknowledgments This work was supported in part by a Grant-in-Aid for Cancer Research (16–6) from the Ministry of Health, Labor, and Welfare, and by a Grant-in-Aid from the Foundation for the Promotion of Defense Medicine.

References

1. Goldhirsch A, Wood W, Gelber R, Coates A, Thurlimann B, Senn HJ. Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. Ann Oncol. 2007;18:1133–44.
2. Goldhirsch A, Glick JH, Gelber RD, Coats AS, Thurlimann B, Senn HJ. Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. Ann Oncol. 2005;16:1569–83.
3. Carlson RW, Brown E, Burstein HJ, Gradishar WJ, Hudis CA, Loprinzi C, Mamounas EP, Perez EA, Pritchard K, Ravdin P, Recht A, Somlo G, Theriault RL, Winer EP, Wolff AC. National Comprehensive Cancer Network: NCCN Task Force Report: Adjuvant Therapy for Breast Cancer. J Natl Compr Canc Netw. 2006;4[Suppl 1]:S1-26.
4. Sobin LH, Wittekind C, editors. TNM Classification of malignant tumours. 6th edn. New York: Wiley; 2002.
5. The Japanese Breast Cancer Society. General rules for clinical and pathological recording of breast cancer, 15th edn. Tokyo: Kanehara Shuppan; 2004.
6. Fukutomi T. Manual for management of breast cancer (in Japanese). Tokyo: Medical View; 1996.
7. Lyman GH, Giuliano AE, Somerfield MR, Benson AB 3rd, Bodurka DC, Burstein HJ, Cochran AJ, Cody HS 3rd, Edge SB, Galper S, Hayman JA, Kim TY, Perkins CL, Podoloff DA, Sivasubramaniam VH, Turner RR, Wahl R, Weaver DL, Wolff AC, Winer EP. American Society of Clinical Oncology guideline recommendations for sentinel lymph node biopsy in early-stage breast cancer. J Clin Oncol. 2005;23:7703–20.
8. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomized trials. Lancet. 1998;351:1451–87.
9. Watanabe T, Sonoo H. Endocrine options for breast cancer treatment: looking beyond tamoxifen. Breast Cancer. 2000;7:345–9.
10. Umemura S, Kurosumi M, Moriya T, Oyama T, Arihiro K, Yamashita H, Umekita Y, Komoike Y, Shimizu C, Fukushima H, Kajiwara H, Akiyama F. Immunohistochemical evaluation for hormone receptors in breast cancer: a practically useful evaluation system and handling protocol. Breast Cancer. 2006;13:232–5.
11. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol. 1998;11:155–68.
12. Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol. 1999;17:1474–81.
13. Ellis MJ, Coop A, Singh B, Mauriac L, Llombert-Cussac A, Janicke F, Miller WR, Evans DB, Dugan M, Brady C, Quebe-Fehling E, Borgs M. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol. 2001;19:3808–16.
14. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19:403–10.
15. Tsuda H, Akiyama F, Kurosumi M, Sakamoto G, Watanabe T. Establishment of histological criteria for high-risk node-negative breast carcinoma for a multi-institutional randomized clinical trial of adjuvant therapy. Japan National Surgical Adjuvant Study of Breast Cancer (NSAS-BC) Pathology Section. Jpn J Clin Oncol. 1998;28:486–91.
16. Kouno T, Shimizu C, Watanabe T, Tsuda H, Akiyama F, Kurosumi M, Sakamoto G. A reliable nuclear grading system for primary breast cancer for selecting high risk invasive ductal carcinoma among node negative patients. Proc Am Soc Clin Oncol. 2003;39:113.
17. Tsuda H. HER-2 (c-erbB-2) test update: present status and problems. Breast Cancer. 2006;13:236–48.
18. Klapper LN, Kirschbaum MH, Sela M, Yarden Y. Biochemical and clinical implications of the ErbB/HER signaling network of growth factor receptors. Adv Cancer Res. 2000;77:25–79.
19. Ellis IO, Bartlett J, Dowsett M, Humphreys S, Jasani B, Miller K, Pinder SE, Rhodes A, Walker R. Updated recommendations for HER2 testing in the UK. J Clin Pathol. 2004;57:233–7.
20. Carlson RW, Moench SJ, Hammond ME, Perez EA, Burstein HJ, Allred DC, Vogel CL, Goldstein LJ, Somlo G, Gradishar WJ, Hudis CA, Jahanzeb M, Stark A, Wolff AC, Press MF, Winer EP, Paik S, Ljung BM. NCCN HER2 Testing in Breast Cancer Task Force: HER2 testing in breast cancer: NCCN Task Force report and recommendations. J Natl Compr Canc Netw. 2006;4[Suppl 3]:S1–22.
21. Ross JS, Fletcher JA. HER-2/neu (c-erb-B2) gene and protein in breast cancer. Am J Clin Pathol. 1999;112:S53–67.
22. Paik S, Bryant J, Tan-Chiu E, Yothers G, Park C, Wickerham DL, Wolmark N. HER2 and choice of adjuvant chemotherapy for invasive breast cancer: National Surgical Adjuvant Breast and Bowel Project Protocol B-15. J Natl Cancer Inst. 2000;92:1991–8.
23. Toi M, Nakamura S, Kuroi K, Iwata H, Ohno S, Masuda N, Kusama M, Yamazaki K, Hisamatsu K, Sato Y, Kashiwaba M, Kaise H, Kurosumi M, Tsuda H, Akiyama F, Ohashi Y, Takatsuka Y. For Japan Breast Cancer Research Group (JBCRG): Phase II study of preoperative sequential FEC and docetaxel predicts of pathological response and disease free survival. Breast Cancer Res Treat. 2007. doi:10.1007/s10549-007-9744-z.
24. Hayes DF, Thor AD, Dressler LG, Weaver D, Edgerton S, Cowan D, Broadwater G, Goldstein LJ, Martino S, Ingle JN, Henderson IC, Norton L, Winer EP, Hudis CA, Ellis MJ, Berry DA. Cancer, Leukemia Group B (CALGB) Investigators: HER2 and response

to paclitaxel in node-positive breast cancer. N Engl J Med. 2007;357:1496–506.
25. Yoshimoto M. Time-dependent interrelationships between pathological peognostic factors, relapse rate in breast cancer patients (in Japanese). Nippon Geka Gakkai Zasshi. 1993;94:1131–43.
26. Vleugel MM, Bos R, van der Groep P, Greijer AE, Shvarts A, Stel HV, van der Wall E, van Diest PJ. Lack of lymphangiogenesis during breast carcinogenesis. J Clin Pathol. 2004;57:746–51.
27. Mohammed RA, Green A, El-Shikh S, Paish EC, Ellis IO, Martin SG. Prognostic significance of vascular endothelial cell growth factors -A, -C and -D in breast cancer and their relationship with angio- and lymphangiogenesis. Br J Cancer. 2007;96:1092–100.
28. Arnaout-Alkarain A, Kahn HJ, Narod SA, Sun PA, Marks AN. Significance of lymph vessel invasion identified by the endothelial lymphatic marker D2–40 in node negative breast cancer. Mod Pathol. 2007;20:183–91.
29. Mook S, Van't Veer LJ, Rutgers EJ, Piccart-Gebhart MJ, Cardoso F. Individualization of therapy using Mammaprint: from development to the MINDACT Trial. Cancer Genomics Proteomics. 2007;4:147–55.
30. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE Jr, Wickerham DL, Wolmark N. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol. 2006;24:3726–34.
31. Kaklamani V. A genetic signature can predict prognosis and response to therapy in breast cancer: Oncotype DX. Expert Rev Mol Diagn. 2006;6:803–9.
32. Inazawa J, Inoue J, Imoto I. Comparative genomic hybridization (CGH)-arrays pave the way for identification of novel cancer-related genes. Cancer Sci. 2004;95:559–63.
33. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001;98:10869–74.
34. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci USA. 2003;100:8418–23.
35. Turner NC, Reis-Filho JS. Basal-like breast cancer and the BRCA1 phenotype. Oncogene. 2006;25:5846–53.
36. Tsuda H, Takarabe T, Hasegawa F, Fukutomi T, Hirohashi S. Large, central acellular zones indicating myoepithelial tumor differentiation in high-grade invasive ductal carcinomas as markers of predisposition to lung and brain metastases. Am J Surg Pathol. 2000;24:197–202.
37. Tsuda H, Tani Y, Weisenberger J, Kitada S, Hasegawa T, Murata T, Tamai S, Hirohashi S, Matsubara O, Natori T. Frequent KIT and epidermal growth factor receptor overexpressions in undifferentiated-type breast carcinomas with 'stem-cell-like' features. Cancer Sci. 2005;96:333–9.
38. Rakha EA, El-Sayed ME, Green AR, Paish EC, Lee AH, Ellis IO. Breast carcinoma with basal differentiation: a proposal for pathology definition based on basal cytokeratin expression. Histopathology. 2007;50:434–8.
39. Tsuda H, Takarabe T, Akashi-Tanaka S, Fukutomi T, Nanasawa T, Watanabe T. Evaluation of histopathological criteria for identifying node-negative breast cancer with high risk of early recurrence in the NSAS-BC protocol study. Breast Cancer. 2000;7:201–9.
40. Fulford LG, Reis-Filho JS, Ryder K, Jones C, Gillett CE, Hanby A, Easton D, Lakhani SR. Basal-like grade III invasive ductal carcinoma of the breast: patterns of metastasis and long-term survival. Breast Cancer Res. 2007;9:R4.

CME

What Causes Discrepancies in HER2 Testing for Breast Cancer?

A Japanese Ring Study in Conjunction With the Global Standard

Shinobu Umemura, MD, PhD,[1] R. Yoshiyuki Osamura, MD, PhD,[1] Futoshi Akiyama, MD, PhD,[2] Keiichi Honma, MD, PhD,[3] Masafumi Kurosumi, MD, PhD,[4] Hironobu Sasano, MD, PhD,[5] Satoshi Toyoshima, MD, PhD,[6] Hitoshi Tsuda, MD, PhD,[7] Josef Rüschoff, MD,[8] and Goi Sakamoto, MD, PhD[9]

Key Words: Breast cancer; HER2 testing; Assessment; Ring study; Multicenter study; Immunohistochemistry; Fluorescence in situ hybridization

DOI: 10.1309/AJCP5UUMFMA5ZKII

Upon completion of this activity you will be able to:
- apply the ASCO/CAP guidelines for HER2 testing of breast carcinomas.
- define the immunohistochemical reactivity rate for HER2 in breast carcinoma that should trigger additional testing by FISH.
- analyze causes of discrepancy in HER2 testing of breast carcinomas by IHC.
- outline a protocol for single institutional guidelines for combined use of IHC and FISH for HER2 testing in breast carcinomas, incorporating quality assurance considerations.

The ASCP is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The ASCP designates this educational activity for a maximum of 1 AMA PRA Category 1 Credit ™ per article.

The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Questions appear on p 1006. Exam is located at www.ascp.org/ajcpcme.

Abstract

We assessed interinstitutional and interobserver consistency of human epidermal growth factor receptor type-2 (HER2) testing using immunohistochemical analysis and fluorescence in situ hybridization (FISH) in a set of 20 breast cancer samples among 10 institutions in Japan and a Herceptin adjuvant study participating laboratory in Germany and identified factors that may lead to discordant results.

We found a good agreement in immunohistochemical HER2 scoring between the coordinating institution and 10 participating laboratories ($\kappa = 0.718$) and excellent agreement for FISH ($\kappa = 0.900$). The results of a comparison between 10 Japanese laboratories and the German laboratory was good for immunohistochemical studies ($\kappa = 0.713$) and excellent for FISH ($\kappa = 0.887$). FISH retesting of equivocal samples (2+ immunohistochemically) improved agreement. Discrepancies between results were attributed to the evaluation process in 33.0% of the samples, staining procedures in 25.0%, and a combination of the two in 41.7%. Evaluation of samples according to the American Society of Clinical Oncology/College of American Pathologists guideline increased the number of 2+ immunohistochemical scores. By performing FISH retesting for these samples, consistency among multiple institutions could be archived. The quality of the staining procedures performed and the consistency of evaluations require regular assessment.

The human epidermal growth factor receptor type-2 (*HER2*) gene encodes a protein (185 kDa) that is a cell surface receptor with tyrosine kinase activity.[1] Amplification of the *HER2* gene and/or overexpression of its protein product has been shown in 25% to 30% of breast cancers.[2,3] Moreover, HER2 status is an important factor in predicting prognosis[2,4] and selection of systemic therapies for treatment.[5-9] Overall, *HER2* gene amplification is associated with a poor clinical outcome,[2,4] and, accordingly, HER2 status has been added to the risk category of the St Gallen consensus recommendation.[10] Overexpression of HER2 protein is also associated with resistance to endocrine therapy that may be specific to selective estrogen receptor modulator therapies, such as tamoxifen, but not to aromatase inhibitors.[5,6,11] In contrast, HER2 overexpression correlates with a response to treatment with anthracyclines and taxanes.[7-9]

Accumulating evidence indicates that trastuzumab (Herceptin) is effective not only for the treatment of metastatic breast cancer but also for early breast cancer with HER2 overexpression. International clinical trials[12-16] have revealed that trastuzumab treatment for primary breast cancer in the adjuvant setting reduced the risk of recurrence and mortality. Based on the results of these trials, trastuzumab has been included in the National Comprehensive Cancer Network guidelines and the St Gallen consensus recommendations. In 2006, the European Medicine Agency and the US Food and Drug Administration approved trastuzumab for primary breast cancer in the adjuvant setting. With these approvals, an increased number of patients may be able to receive treatment with trastuzumab based on HER2 testing results.

Misdiagnosis of HER2 overexpression can result in the loss of opportunity for patients to receive the benefits of trastuzumab treatment or in patients being overtreated. Therefore, accuracy in HER2 testing is of significant clinical benefit.

It has been reported that the efficacy of trastuzumab depends on the extent of HER2 overexpression: A tumor with a 2+ immunohistochemical score has a response rate of 0%, whereas a tumor with a 3+ immunohistochemical score has a response rate of 35%.[17] Although the significance of accurate HER2 testing has been emphasized, HER2 testing is not subject to external quality assurance in all countries, despite the fact that evaluation of HER2 serves as a major conclusive factor in the decision to treat with trastuzumab.

HER2 gene amplification was first examined by using Southern blotting in the early phase of a clinical study by Slamon et al.[2] An alternative method for HER2 detection is fluorescence in situ hybridization (FISH). In the early phase of a validation study, HER2 protein overexpression was examined immunohistochemically using anti-HER2 monoclonal antibodies 4D5 and CB11 (denoted the Clinical Trial Assay).[16] HER2 detection using a polyclonal antibody can be more sensitive but is less specific than using a monoclonal antibody. Although immunohistochemical analysis is now relatively inexpensive and universally available in research laboratories, it does not produce results as reliably consistent as those observed with FISH. In 3 clinical studies, the population of patients with tumors categorized as 2+ immunohistochemically varied from 12.7% to 39.5%, and the rate of *HER2* gene amplification in tumors scored 2+ immunohistochemically varied from 17.9% to 48.1%.[16,18,19] Although these data may represent a significant diversity in breast cancer tissue samples with regard to HER2, it is important to consider the sensitivity and specificity of immunohistochemical and FISH analyses.

No assessment system for the standardization of immunohistochemical or FISH analysis of HER2 has been established in Japan thus far. However, in other countries, a standardization process is in place, ie, the Nordic Immunohistochemical Quality Control (http://www.nordiqc.org/news.htm), United Kingdom National External Quality Assessment Service (http://www.ukneqas.org.uk/), the College of American Pathologists (CAP; http://www.cap.org/apps/cap.portal), and the Royal College of Pathologists of Australia Quality Assurance Program (http://www.rcpaqapa.netcore.com.au/index.html).

To investigate the consistency of HER2 testing in Europe, Dowsett et al[20] conducted an international ring study with 5 pathologists, each from a different country, applying immunohistochemical analysis and FISH to 20 slide sets. We conducted a Japanese ring study with 10 participating laboratories, responsible for diagnosing approximately 80% of breast cancer samples in Japan, and 1 laboratory in Germany that participated in the Herceptin adjuvant (HERA) trial.

Materials and Methods

Participants

To compare and assess immunohistochemical and FISH analyses of HER2 expression by different laboratories, 7 institutions in Japan (Tokai University School of Medicine, Isehara; The Cancer Institute Japanese Foundation for Cancer Research, Koto; Niigata Cancer Center Hospital, Niigata; the Saitama Cancer Center, Kita-adachi; Tohoku University School of Medicine, Sendai; Kitakyushu Municipal Medical Center, Kitakyushu; and the National Defense Medical College, Tokorozawa), 3 commercial laboratories (SRL, Tachikawa; BML, Kawagoe; and Mitsubishi Chemical Medience, Itabashi), and 1 laboratory that was a site for a HERA trial (Institut für Pathologie, Klinikum Kassel, Targos Molecular Pathology, Kassel, Germany) participated in this ring study **Figure 1**. The study was coordinated by the Tokai University School of Medicine and approved by the institutional review board of Tokai University.

Sample Selection and Distribution

We selected 20 cases of invasive breast cancer from the surgical pathology files of Tokai University Hospital. The breast cancer tissue samples had previously been tested and were selected to represent a relatively higher proportion of equivocal cases for the purpose of assessment. All of the specimens had been fixed with formalin (12-48 hours) and embedded in paraffin blocks. Tissue sections,

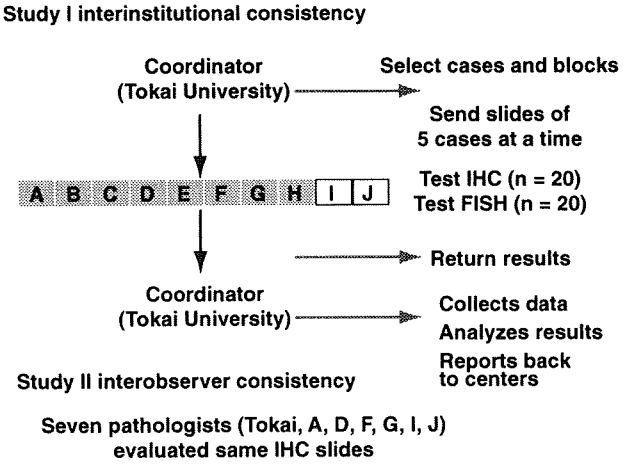

Figure 1 Study designs for studies I and II. Study I examined interinstitutional consistency, and study II examined interobserver consistency of HER2 testing. Institutions A-H participated in immunohistochemical analysis (IHC) and fluorescence in situ hybridization (FISH) analysis, and institutions I and J evaluated IHC results.

DOI: 10.1309/AJCP5UUMFMA5ZKII

4 to 6 μm thick, were mounted on silane-coated slides. A set of 5 cases, with 2 unstained slides for each case, was sent every 2 weeks to participants for immunohistochemical and FISH analyses. The slides for FISH analysis were sent after receiving the results of the immunohistochemical analysis to avoid bias from the FISH results. In this way, an identical series of 20 cases was evaluated independently for immunohistochemical and FISH detection of HER2 expression. Of the 10 participating, 8 institutions performed HER2 testing by immunohistochemical analysis (HercepTest, DakoCytomation, Carpinteria, CA) and FISH (PathVysion, Vysis, Downers Grove, IL) analyses; 2 participants performed immunohistochemical analysis only.

Study Design and Data Analysis

This study was designed to examine interinstitutional consistency (study I) and interobserver consistency (study II) in the analysis of tissue samples. Sample selection and distribution of sections for study I was described in the preceding section, with the evaluated results analyzed by the study coordinator. For study II, the goal was to examine interobserver consistency. Seven pathologists each evaluated one set of 20 cases that were stained at the Tokai University School of Medicine, the same set of 20 cases evaluated in study I. To evaluate the significance of the assessment system proposed by the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guidelines for interobserver consistency in immunohistochemical analysis, evaluation systems described in the manufacturer's protocol (study IIA) and those of the ASCO/CAP guideline were compared (study IIB).

The results were estimated by concordance rate and κ value. A κ value of more than 0.75 represented excellent agreement, values from 0.4 to 0.75 represented fair to good agreement, and values less than 0.4 represented poor agreement beyond chance. Concordance rates between the coordinating laboratory and participating Japanese laboratories and between the HERA laboratory and the Japanese laboratories were evaluated.

Immunohistochemical Analysis and FISH

Immunohistochemical detection kits for HER2 containing the same lot of polyclonal antibody (HercepTest) were distributed to all participating laboratories. Sections were stained according to the manufacturer's protocols. Staining results were evaluated using the criteria 0, 1+, 2+, and 3+ according to the HercepTest kit instructions, which were the standardized criteria at the time, for studies I and IIA **Image 1**. In study IIB, the ASCO/CAP guidelines were used as the staining criteria.[21]

For FISH detection of HER2, HER2/CEP17 probe kits (Vysis) were distributed and used according to the manufacturer's protocol. Signal numbers for the *HER2* gene (labeled with SpectrumOrange, Vysis) and the *CEP17* gene (labeled with SpectrumGreen, Vysis) were counted in more than 20 tumor cells from each site, and the ratio of the HER2/CEP17 signal numbers was calculated. The results were interpreted as positive when the signal ratio of HER2/CEP17 was equal or greater than 2.0 and negative when it was less than 2.0 according to the manufacturer's protocol.

Results

Interinstitutional Consistency

The results of study I are shown in **Table 1** and **Table 2**. Of 20 samples analyzed immunohistochemically, 14 (70%) tumors were grouped in the same category when the results were categorized into 2 groups as 0, 1+/2+ and 3+; and 8 (40%) of 20 scores were consistent when the results were categorized as negative (0, 1+), equivocal (2+), and positive (3+) (Table 1). The recorded immunohistochemical results were in good agreement between the participating Japanese laboratories and the coordinating laboratory (κ = 0.718), and agreement was also good between participants and the HERA laboratory (institution E in Table 1) (κ = 0.713). For FISH analyses, results for 17 (85%) of 20 samples were consistent for all participants (Table 2). Discrepancies in results were mainly observed for samples with a HER2/CEP17 signal ratio close to 2.0. The FISH results were in excellent agreement between participants and the coordinator (κ = 0.900) and between participants and the HERA laboratory (institution E in Table 2) (κ = 0.887).

FISH Retesting for Cases Immunohistochemically Scored as 2+ Increased Agreement

To verify the algorithm for "Indication of Trastuzumab and HER2 Testing,"[22] simulation analyses were performed. Initially, analysis of study I was conducted to determine whether retesting of FISH for cases scored immunohistochemically as 2+ would improve agreement. There were 11 cases that were determined to be 2+ by at least 1 institution. When these cases were retested by FISH, agreement in scoring improved in 8 of the 11 cases. In the remaining 3 cases, scores were lowered or unchanged **Table 3**. A second analysis was used to validate whether the 2 trees of algorithms are acceptable from the perspective of clinical benefit. According to the distribution of results organized into the FISH tree **Figure 2**, 46.7% of cases had indications for trastuzumab treatment. Based on the distribution of results into the immunohistochemical analysis tree (Figure 2), 45.6% had indications for trastuzumab treatment. Retesting samples by FISH closed the gap further, demonstrating that

DOI: 10.1309/AJCP5UUMFMA5ZKII

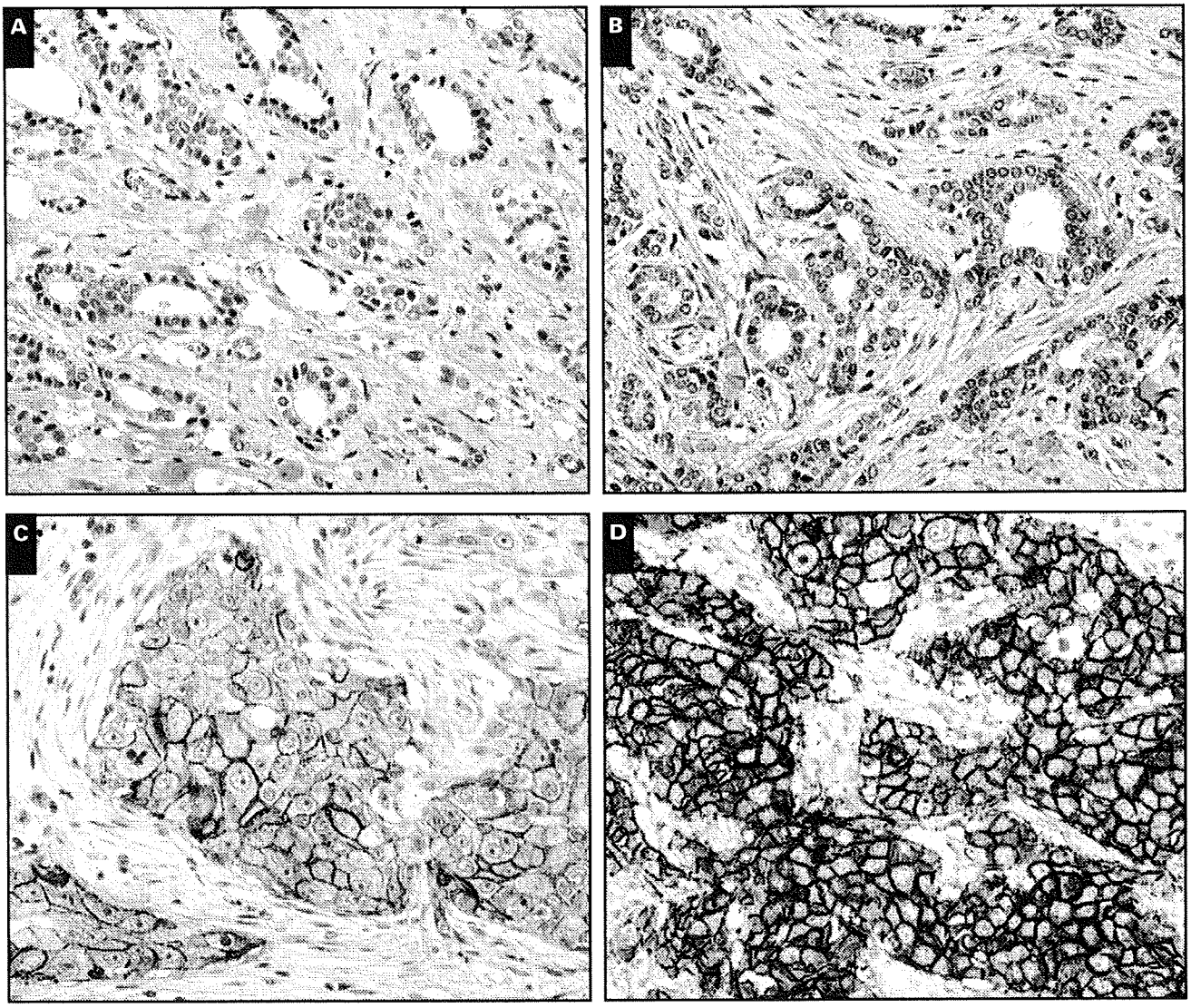

Image 1 Scoring criteria for immunohistochemical examination. **A**, Carcinoma cells lack positive reactivity for HER2 protein (score 0; ×100). **B**, A weakly positive reaction is found, but they are not completely bound to the cell membrane (score 1+; ×100). **C**, Weak to moderate, complete membrane staining is detected in >10% of tumor cells (score 2+; ×100). **D**, Strong, complete membrane staining is observed in >10% of tumor cells (score 3+ in studies I and IIA) and in >30% of tumor cells (score 3+ in study IIB) (×100).

retesting by using FISH can minimize the prescription of treatment that will have no effect or too much of an effect, depending on the type of cancer present.

Interobserver Discrepancies and Their Causes

To analyze the cause of discrepancies that occur in evaluating the pathology of tissue samples, interobserver consistency was examined in study II. By using the evaluations provided by each of the 7 pathologists on the Japanese Pathology Board for Optimal Use of Trastuzumab for 1 set of the 20 cases, in which the variable factor of staining procedures had been excluded, we examined interobserver discrepancy. Nine of the evaluations were inconsistent, with discrepancies between immunohistochemical scores of 2+ and 3+ in 2 samples (cases 4 and 6) **Image 2** and between immunohistochemical scores of 2+ and 1+ in 7 samples (cases 3, 5, 7, 9, 10, 18, and 15; study IIA) **Table 4**. The analysis of discrepancies from study I suggested that a complexity of factors, including interobserver diversity and staining procedures (Table 4, study I), accounted for the differences. However, of the 12 samples in which the results were not consistent in study I, 3 (25%) showed complete agreement in study IIA, and interinstitutional concordance was lower in study I. We hypothesize that

DOI: 10.1309/AJCP5UUMFMA5ZKII

Table 1
Analysis of Interinstitutional Concordance by Immunohistochemical Scoring Results*

Case No.	Coordinator	Institution										Concordance (%)
		A	B	C	D	E	F	G	H	I	J	
11	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	100
12	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	100
13	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	100
19	3+	3+	3+	3+	3+	3+	2+	3+	3+	3+	3+	91
20	3+	3+	3+	3+	3+	3+	2+	3+	3+	3+	3+	91
4	3+	2+	3+	2+	3+	2+	2+	3+	2+	2+	2+	64
8	3+	3+	3+	3+	3+	3+	3+	3+	3+	3+	2+	91
6	2+	3+	3+	3+	2+	3+	2+	2+	2+	2+	2+	64
5	2+	2+	2+	3+	3+	2+	2+	2+	2+	2+	2+	82
7	2+	2+	1+	2+	2+	2+	2+	1+	2+	2+	1+	73
9	2+	2+	2+	2+	2+	2+	2+	1+	2+	2+	1+	82
10	2+	1+	1+	2+	0	2+	1+	1+	2+	1+	1+	64
18	1+	2+	1+	2+	2+	2+	1+	1+	1+	2+	1+	55
15	1+	1+	1+	2+	1+	1+	1+	1+	2+	2+	1+	73
3	1+	1+	1+	1+	1+	1+	1+	1+	1+	1+	0	100
16	1+	1+	1+	2+	0	1+	0	1+	1+	1+	0	91
2	1+	1+	1+	0	1+	1+	0	1+	1+	0	0	100
14	1+	1+	1+	1+	0	1+	0	0	0	0	0	100
1	0	0	1+	0	0	0	1+	1+	0	0	0	100
17	0	0	0	0	0	1+	0	0	0	0	0	100

* 0 and 1+ were considered negative; 2+, equivocal; and 3+, positive.

Table 2
Analysis of Interinstitutional Concordance by Fluorescence In Situ Hybridization Results*

Case No.	Coordinator	Institution								Concordance (%)
		A	B	C	D	E	F	G	H	
11	7.1	7.4	3.8	18.7	8.3	4.9	10.0	6.2	7.3	100
12	4.6	9.4	4.5	13.9	4.4	4.6	7.8	4.2	6.8	100
13	7.3	9.8	9.6	14.1	8.4	7.4	10.3	4.2	5.2	100
19	6.7	6.3	8.1	11.5	3.9	5.5	8.0	13.3	6.2	100
20	6.6	7.7	5.9	7.4	5.4	5.6	6.7	4.5	4.5	100
4	6.1	8.9	4.7	10.7	5.6	5.2	7.9	3.9	5.3	100
8	3.2	4.9	2.7	12.5	2.4	5.0	4.4	5.4	2.7	100
6	6.5	4.4	4.4	20.9	5.4	3.9	4.6	3.9	2.7	100
5	2.5	4.1	3.5	24.7	2.5	2.6	1.6	1.6	1.9	67
7	2.5	2.2	2.1	1.8	1.9	1.3	2.6	3.0	1.8	56
9	1.3	1.2	1.3	1.5	1.5	1.0	1.5	1.2	1.5	100
10	1.6	1.2	1.4	2.0	1.2	1.0	1.0	1.4	1.3	89
18	1.2	1.2	1.2	1.0	1.2	1.2	1.7	1.0	1.1	100
15	1.1	1.3	1.3	1.0	1.0	0.9	1.4	1.4	1.0	100
3	1.3	0.9	1.2	1.0	1.3	1.0	1.4	1.5	1.2	100
16	1.4	1.2	1.0	1.1	1.2	1.1	1.2	1.5	1.1	100
2	1.8	1.3	1.0	1.0	1.5	1.1	1.1	1.0	1.2	100
14	1.9	1.4	1.1	1.1	1.0	1.2	1.3	1.1	1.1	100
1	1.2	1.4	1.2	1.0	1.4	1.1	1.9	1.2	0.9	100
17	1.1	1.2	1.1	1.0	1.0	1.1	1.1	1.1	0.9	100

* Data are given as the HER2/CEP17 ratio. A ratio <2.0 was considered negative, and a ratio ≥2.0, positive.

staining procedures were the cause. In 4 of 12 cases, the interobserver discrepancy present in study IIA was identical to the concordance found in study I. For these cases, interobserver discrepancy was considered to be the cause. In the other 5 of 12 cases, both interobserver and interinstitutional discrepancies suggest that staining procedures and interobserver discrepancies are possible causes.

Significance of ASCO/CAP Guideline

To study the significance of the ASCO/CAP interpretive criteria, we conducted study IIB. The variable factor is the only evaluation system in study II; therefore, we can analyze the significance of the proposed criteria in comparison with previous criteria and the ASCO/CAP guidelines while excluding the influences of staining procedures. Revised

Table 3
Analysis of FISH Retesting for Samples With Immunohistochemical Scores of 2+*

			Institution											
	Coordinator		A		B		C		D		E		F	
Case No.	IHC	FISH	IHC	FISH	IHC	FISH	IHC	FISH	IHC	FISH	IHC	FISH	IHC	FISH
11	3+	7.1	3+	7.4	3+	3.8	3+	18.7	3+	8.3	3+	4.9	3+	10.0
12	3+	4.6	3+	9.4	3+	4.5	3+	13.9	3+	4.4	3+	4.6	3+	7.8
13	3+	7.3	3+	9.8	3+	9.6	3+	14.1	3+	8.4	3+	7.4	3+	10.3
19	3+	6.7	3+	6.3	3+	8.1	3+	11.5	3+	3.9	3+	5.5	2+	8.0
20	3+	6.6	3+	7.7	3+	5.9	3+	7.4	3+	5.4	3+	5.6	2+	6.7
4	3+	6.1	2+	8.9	3+	4.7	2+	10.7	3+	5.6	2+	5.2	2+	7.9
8	3+	3.2	3+	4.9	3+	2.7	3+	12.5	3+	2.4	3+	5.0	3+	4.4
6	2+	6.5	3+	4.4	3+	4.4	3+	20.9	2+	5.4	3+	3.9	2+	4.6
5	2+	2.5	2+	4.1	2+	3.5	3+	24.7	3+	2.5	2+	2.6	2+	1.6
7	2+	2.5	2+	2.2	1+	2.1	2+	1.8	2+	1.9	2+	1.3	2+	2.6
9	2+	1.3	2+	1.2	2+	1.3	2+	1.5	2+	1.5	2+	1.0	2+	1.5
10	2+	1.6	1+	1.2	1+	1.4	2+	2.0	0	1.2	2+	1.0	1+	1.0
18	1+	1.2	2+	1.2	1+	1.2	2+	1.0	2+	1.2	2+	1.2	1+	1.7
15	1+	1.1	1+	1.3	1+	1.3	2+	1.0	1+	1.0	1+	0.9	1+	1.4
3	1+	1.3	1+	0.9	1+	1.2	1+	1.0	1+	1.3	1+	1.0	1+	1.4
16	1+	1.4	1+	1.2	1+	1.0	2+	1.1	0	1.2	1+	1.1	0	1.2
2	1+	1.8	1+	1.3	1+	1.0	0	1.0	1+	1.5	1+	1.1	0	1.1
14	1+	1.9	1+	1.4	1+	1.1	1+	1.1	0	1.0	1+	1.2	0	1.3
1	0	1.2	0	1.4	1+	1.2	0	1.0	0	1.4	0	1.1	1+	1.9
17	0	1.1	0	1.2	0	1.1	0	1.0	0	1.0	1+	1.1	0	1.1

FISH, fluorescence in situ hybridization; IHC, immunohistochemical analysis.
* Data are given for IHC as immunohistochemical scores (0 and 1+, negative; 2+, equivocal; 3+, positive) and for FISH as the HER2/CEP17 ratio (<2.0, negative; ≥2.0, positive).

assessments by at least 1 pathologist were detected in 3 cases **Table 5**. As a result, concordance between sample designations increased, decreased, and remained unchanged for the 3 cases, respectively. It is noteworthy that most of the changes were from an immunohistochemical score of 3+ to 2+.

The FISH results were reviewed according to the ASCO/CAP criteria, and interpretations of equivocal (1.8-2.2) were frequently identified in discrepant cases according to manufacturer's criteria (cases 5 and 7).

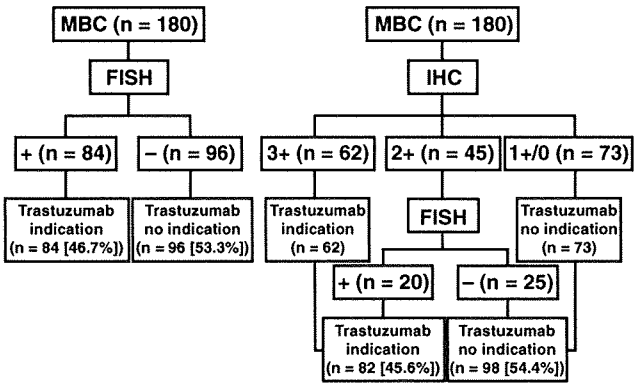

Figure 2 Simulation analysis for retesting fluorescence in situ hybridization (FISH) samples for immunohistochemical analysis (IHC) scores of 2+ according to the HER2 testing algorithm. MBC, metastatic breast cancer.

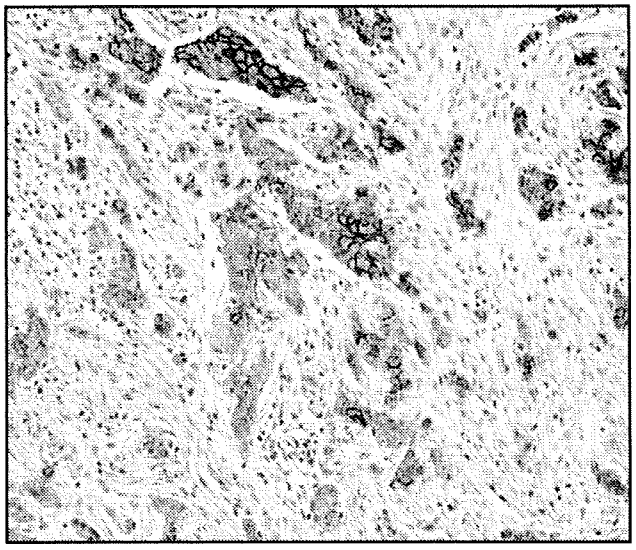

Image 2 (Case 4) Immunohistochemical features. Tumor cells with varied intensity for HER2 protein are heterogeneously distributed (×100).

DOI: 10.1309/AJCP5UUMFMA5ZKII

Institution					
G		H		Concordance (%)	
IHC	FISH	IHC	FISH	IHC	ICH2+/FISH
3+	6.2	3+	7.3	100	
3+	4.2	3+	6.8	100	
3+	4.2	3+	5.2	100	
3+	13.3	3+	6.2	89	100
3+	4.5	3+	4.5	89	100
3+	3.9	2+	5.3	56	100
3+	5.4	3+	2.7	100	
2+	3.9	2+	2.7	56	100
2+	1.6	2+	1.9	78	67
1+	3.0	2+	1.8	78	67
1+	1.2	2+	1.5	89	100
1+	1.4	2+	1.3	56	89
1+	1.0	1+	1.1	56	100
1+	1.4	2+	1.0	78	
1+	1.5	1+	1.2	100	89
1+	1.5	1+	1.1	89	100
1+	1.0	1+	1.2	100	
0	1.1	0	1.1	100	
1+	1.2	0	0.9	100	
0	1.1	0	0.9	100	

Discussion

Accuracy in HER2 testing is very important for the treatment of patients. Large clinical trials such as the North Central Cancer Treatment Group, National Surgical Adjuvant Breast and Bowel Project, and HERA require a standardization of HER2 testing, thus emphasizing global quality control. The present Japanese ring study demonstrates good agreement for immunohistochemical detection of HER2 and excellent agreement for HER2 detection using FISH despite a higher proportion of equivocal ratings (2+). Agreement levels between participants and the coordinator (κ values for immunohistochemical analysis and FISH of 0.718 and 0.900, respectively) and the HERA laboratory (κ values for immunohistochemical analysis and FISH of 0.713 and 0.887, respectively) were almost identical.

In the present study, an attempt was made to exclude variable factors in the technical procedures by use of a detection kit containing the same lot of antibody for immunohistochemical analysis and FISH probes and also the use of the same protocols. However, the tissue processing (eg, fixation of tumor samples, absorbance, tissue embedding) before immunohistochemical analysis was conducted in different laboratories in different countries and was not controlled. Therefore, tissue preparation variables, if present, were maintained in this study. Irrespective of these variable regional factors, however, agreement was obtained during analysis of pathology between laboratories in Japan and Germany. In a previous study, Dowsett et al[20] conducted an international ring study with 5 participants from different countries—the Netherlands, Canada, France, Belgium, and Germany—using 20 sets for immunohistochemical analysis and FISH. The

Table 4
Interobserver Discrepancies in Immunohistochemical Results and Causes of Discrepant Results*

	Study I								Study II							
		Institution								Institution						
Case No.	Coordinator	A	D	F	G	I	J	Concordance (%)	Coordinator	A	D	F	G	I	J	Concordance (%)
11	3+	3+	3+	3+	3+	3+	3+	100	3+	3+	3+	3+	3+	3+	3+	100
12	3+	3+	3+	3+	3+	3+	3+	100	3+	3+	3+	3+	3+	3+	3+	100
13	3+	3+	3+	3+	3+	3+	3+	100	3+	3+	3+	3+	3+	3+	3+	100
19	3+	3+	3+	2+	3+	3+	3+	86†	3+	3+	3+	3+	3+	3+	3+	100†
20	3+	3+	3+	2+	3+	3+	3+	86†	3+	3+	3+	3+	3+	3+	3+	100†
4	3+	2+	3+	2+	3+	2+	2+	57‡	3+	2+	2+	3+	2+	2+	3+	57‡
8	3+	3+	3+	3+	3+	3+	2+	86†	3+	3+	3+	3+	3+	3+	3+	100†
6	2+	3+	2+	2+	2+	2+	2+	86‡	2+	2+	2+	2+	2+	2+	3+	86‡
5	2+	2+	3+	2+	2+	2+	2+	86‡	2+	2+	1+	2+	2+	2+	2+	86‡
7	2+	2+	2+	2+	1+	2+	1+	71§	2+	1+	2+	2+	2+	2+	2+	86§
9	2+	2+	2+	2+	1+	2+	1+	71§	2+	2+	2+	2+	1+	2+	2+	86§
10	2+	1+	0	1+	1+	1+	1+	86§	2+	2+	1+	2+	1+	2+	2+	71§
18	1+	2+	2+	1+	1+	2+	1+	57§	1+	1+	1+	1+	1+	1+	2+	86§
15	1+	1+	1+	1+	1+	2+	1+	86‡	1+	2+	1+	1+	1+	1+	1+	86‡
3	1+	1+	1+	1+	1+	1+	0	100§	1+	1+	1+	1+	1+	1+	2+	86§
16	1+	1+	0	0	1+	1+	0	100	1+	1+	1+	1+	0	0	1+	100
2	1+	1+	1+	0	1+	0	0	100	1+	0+	0	0	0	0	0	100
14	1+	1+	0	0	0	0	0	100	1+	1+	1+	1+	0	0	1+	100
1	0	0	0	1+	1+	0	0	100	0	0	0	0	0	0	0	100
17	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	100

* Data are given as immunohistochemical scores (0 and 1+, negative; 2+, equivocal; 3+, positive).
† Complete agreement in interobserver consistency in study II, with lower concordance in study I.
‡ Identical concordance between studies I and II.
§ Different concordance between studies I and II.

DOI: 10.1309/AJCP5UUMFMA5ZKII

Table 5
Significance of Assessment According to the ASCO/CAP Guidelines*

Case No.	Coordinator	Institution						Concordance (%)
		A	D	F	G	I	J	
Study IIA								
4	3+	2+	2+	3+	2+	2+	3+§	57†
8	3+	3+	3+§	3+	3+	3+	3+§	100‡
6	2+	2+	2+	2+	2+	2+	3+§	86
Study IIB								
4	3+	2+	2+	3+	2+	2+	2+§	71†
8	3+	3+	2+§	3+	3+	3+	2+§	71‡
6	2+	2+	2+	3+	2+	2+	2+§	86

ASCO/CAP, American Society of Clinical Oncology/College of American Pathologists.
* For study IIA, immunohistochemical staining results were evaluated by using the criteria 0, 1+, 2+, and 3+ (0 and 1+, negative; 2+, equivocal; 3+, positive) according to the HercepTest kit instructions. In study IIB, the ASCO/CAP guidelines were used as the evaluation criteria. According to the ASCO/CAP guidelines, the concordance rate was increased† or decreased,‡ and the interpretations from 3+ to 2+ are increased.§

concordance rate for immunohistochemical analysis was 45% (9/20) in categories of negative, equivocal, and positive; and for FISH, the rate was 80% (16/20). In our Japanese ring study, despite the increased number of participants (11 including the coordinator), the concordance rate was similar to that in the study by Dowsett et al.[20]

The goal for this study was to identify causes of discrepancies in HER2 detection in breast cancer samples. By using interinstitutional and interobserver conditions for analysis of the same series of tumors, we tried to pinpoint factors that contribute to discrepant results. Interinstitutional discrepancies in immunohistochemical analysis were identified in 6 samples categorized as 2+ and 3+ and also in 5 samples categorized as 2+ and 1+. In both cases, discrepancies were related to technical and evaluation methods. In these 12 samples, interobserver study showed that 3 samples (3/12 [25%]) were 100% consistent when the pathologists evaluated sections stained by the same method. Based on these conclusions, discrepancies in results from the interinstitutional study were assumed to be related to tissue processing and staining procedures. Interobserver diversities were identified in 4 samples (4/12 [33%]), and the percentage of discord in the interinstitutional study was the same. Thus, it was assumed that interobserver diversity was the major cause of subsequent discrepancies. The remaining 5 samples (5/12 [42%]) were discordant owing to complex causes of technology and evaluation because interobserver discrepancy was present, and the interinstitutional concordance rate was lower or higher than the interobserver concordance rate. Table 4 shows that the staining procedure was most frequently identified as the cause of discrepancy between cases scored immunohistochemically as 2+ vs 3+, and that technical methods and interobserver diversity were more frequently identified for differences between cases scored immunohistochemically as 1+ vs 2+.

Discrepancies between immunohistochemical evaluations of 1+ and 2+ are clinically critical; therefore, assessment of both staining procedures and evaluation methods should be well controlled. To our knowledge, this is the first ring study designed to clarify the cause of discrepancies in HER2 analysis by immunohistochemical analysis and FISH by minimizing variable factors.

The ASCO/CAP "Guideline for HER2 Testing in Breast Cancer"[21] recently proposed the category of "equivocal" for tumors identified with an immunohistochemical designation of 2+ and a FISH ratio of 1.8 to 2.2. For these samples, reexamination by FISH is recommended. In addition, the criterion for immunohistochemical results of 3+ was redefined "as uniform intense membrane staining of >30% of invasive tumor cells." The present study IIB clearly showed that the new definition increased the proportion of cases designated immunohistochemically as 2+, which would be subsequently examined by FISH according to the ASCO/CAP guidelines. As shown in the simulation analysis, retesting by FISH of samples scored 2+ immunohistochemically increased the concordance rate. Thus, the currently proposed ASCO/CAP guideline can improve evaluation consistency among multiple institutions and provide more reliable identification of the most appropriate patients for trastuzumab treatment.

We assessed the quality and consistency of HER2 testing performed by laboratories in Japan that are responsible for evaluating approximately 80% of breast cancer tissue samples submitted for pathology studies. We found good to excellent agreement among the participants and in comparison with results from a HERA laboratory in Germany. This is the first ring study to evaluate the causes of discrepancies in analysis of breast cancer pathology with regard to HER2 expression by comparing interinstitutional and interobserver results with an effort to minimize technical variables.

DOI: 10.1309/AJCP5UUMFMA5ZKII

From the [1]Department of Pathology, Tokai University School of Medicine, Isehara, Japan; [2]Division of Pathology, The Cancer Institute Japanese Foundation for Cancer Research, Koto; [3]Division of Pathology, Niigata Cancer Center Hospital, Niigata, Japan; [4]Department of Pathology, Saitama Cancer Center, Kita-adachi, Japan; [5]Department of Pathology, Tohoku University School of Medicine, Sendai, Japan; [6]Division of Pathology, Kitakyushu Municipal Medical Center, Kitakyushu, Japan; [7]Department of Basic Pathology, National Defense Medical College, Tokorozawa, Japan; [8]Department of Pathology, Institute of Pathology Nordhessen, Kassel, Germany; and [9]Department of Pathology, Sakamoto Memorial Clinic Academy of Breast Pathology, Shinjuku, Japan.

Supported by Chugai Pharmaceutical, Chuo, Japan; DakoCytomation, Bunkyo, Japan; Abbott Japan, Minato; and SRL, Tachikawa, Japan.

Address reprint requests to Dr Umemura: Dept of Pathology, Tokai University School of Medicine, Shimokasuya 143, Isehara city, Kanagawa prefecture, Japan 259-1193.

Acknowledgments: We thank Oliver Stoss, PhD, Institut für Pathologie, Klinikum Kassel, Targos Molecular Pathology, Kassel, Germany, for support from the central laboratory of the Herceptin adjuvant trials; Yutaka Tokuda, MD, Department of Surgery, and Yukiko Kirimura, Department of Pathology, Tokai University School of Medicine, for support from the coordinator institution; Toru Watanabe, MD, Hamamatsu Oncology Center, Hamamatsu; and Masakazu Toi, MD, Department of Breast Surgery, Kyoto University Graduate School of Medicine, for their clinical advice.

References

1. Akiyama T, Sudo C, Ogawara H, et al. The product of the human c-erbB-2 gene: a 185-kilodalton glycoprotein with tyrosine kinase activity. *Science.* 1986;232:1644-1646.
2. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/*neu* oncogene. *Science.* 1987;235:177-182.
3. 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.
4. Seshadri R, Firgaira FA, Horsfall DJ, et al. Clinical significance of HER-2/*neu* oncogene amplification in primary breast cancer. The South Australian Breast Cancer Study Group. *J Clin Oncol.* 1993;11:1936-1942.
5. Leitzel K, Teramoto Y, Konrad K, et al. Elevated serum c-erbB-2 antigen levels and decreased response to hormone therapy of breast cancer. *J Clin Oncol.* 1995;13:1129-1135.
6. Yamauchi H, O'Neill A, Gelman R, et al. Prediction of response to antiestrogen therapy in advanced breast cancer patients by pretreatment circulating levels of extracellular domain of the HER-2/c-neu protein. *J Clin Oncol.* 1997;15:2518-2525.
7. Paik S, Bryant J, Park C, et al. erbB-2 and response to doxorubicin in patients with axillary lymph node-positive, hormone receptor-negative breast cancer. *J Natl Cancer Inst.* 1998;90:1361-1370.
8. Thor AD, Berry DA, Budman DR, et al. erbB-2, p53, and efficacy of adjuvant therapy in lymph node-positive breast cancer. *J Natl Cancer Inst.* 1998;90:1346-1360.
9. Konecny GE, Thomssen C, Luck HJ, et al. HER-2/*neu* gene amplification and response to paclitaxel in patients with metastatic breast cancer. *J Natl Cancer Inst.* 2004;96:1141-1151.
10. Goldhirsch A, Coates AS, Gelber RD, et al. First: select the target: better choice of adjuvant treatments for breast cancer patients. *Ann Oncol.* 2006;17:1772-1776.
11. Ellis MJ, Coop A, Singh B, et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for erbB-1- and/or erbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. *J Clin Oncol.* 2001;19:3808-3816.
12. Slamon D, Eiermann W, Robert N, et al. Phase III randomized trial comparing doxorubicin and cyclophosphamide followed by docetaxel (ACT) with doxorubicin and cyclophosphamide followed by docetaxel and trastuzumab (ACTH) 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. Abstract 1.
13. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med.* 2006;354:809-820.
14. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med.* 2005;353:1673-1684.
15. Smith I, Procter M, Gelber RD, et al. 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet.* 2007;369:29-36.
16. Dybdal N, Leiberman G, Anderson S, et al. Determination of *HER2* gene amplification by fluorescence in situ hybridization and concordance with the clinical trials immunohistochemical assay in women with metastatic breast cancer evaluated for treatment with trastuzumab. *Breast Cancer Res Treat.* 2005;93:3-11.
17. 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.
18. 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 centers. *J Pathol.* 2003;199:418-423.
19. Yaziji H, Goldstein LC, Barry TS, et al. HER-2 testing in breast cancer using parallel tissue-based methods. *JAMA.* 2004;291:1972-1977.
20. Dowsett M, Hanna WM, Kockx M, et al. Standardisation of HER2 testing: results of an international proficiency-testing ring study. *Mod Pathol.* 2007:20:584-591.
21. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med.* 2007;131:19-26.
22. Umemura S, Sakamoto G, Sasano H, et al. Evaluation of HER2 status: for the treatment of metastatic breast cancers by humanized anti-HER2 monoclonal antibody (trastuzumab). Pathological Committee for Optimal Use of Trastuzumab. *Breast Cancer.* 2001;8:316-320.

Predictive implications of nucleoside metabolizing enzymes in premenopausal women with node-positive primary breast cancer who were randomly assigned to receive tamoxifen alone or tamoxifen plus tegafur-uracil as adjuvant therapy

MASAKAZU TOI[1], TADASHI IKEDA[2], FUTOSHI AKIYAMA[3], MASAFUMI KUROSUMI[4], HITOSHI TSUDA[5], GOI SAKAMOTO[6] and OSAHIKO ABE[7]

[1]Department of Surgery (Breast Surgery), Graduate School of Medicine, Kyoto University, Kyoto; [2]Department of Surgery, Teikyo University School of Medicine, [3]Department of Breast Pathology, Cancer Institute of Japanese Foundation for Cancer Research, Tokyo; [4]Department of Pathology, Saitama Cancer Center Hospital, Ina-machi, [5]Department of Pathology II, National Defense Medical College, Tokorozawa, Saitama; [6]Sakamoto Memorial Clinic, Academy of Breast Pathology, [7]St. Luke's Hospital, Tokyo, Japan

Received May 7, 2007; Accepted June 29, 2007

Abstract. Recent studies have demonstrated that tegafur-uracil (UFT) is useful for the adjuvant treatment of various types of cancers. To determine whether nucleoside metabolizing enzymes could be used to predict the response to UFT treatment in women with primary breast cancer, we retrospectively analyzed archived tumor tissue samples obtained from the 3rd Adjuvant Chemo-Endocrine Therapy for Breast Cancer (ACETBC) study, in which adjuvant treatment with tamoxifen (TAM) plus UFT for 2 years was compared with TAM alone for 2 years. Samples of tumor tissue were obtained from 192 premenopausal women with node-positive invasive breast cancer. The tissue samples were examined immunohistochemically to study the expression of thymidylate synthase (TS), thymidine phosphorylase (TP), and dihydropyrimidine dehydrogenase (DPD), as well as the expression of HER2 and p53. In patients with TS-positive tumors, the risk of relapse was significantly lower in the tamoxifen plus UFT group than in the tamoxifen alone group. After 2 years, however, there was a trend towards a decrease in the relative predictive value (RPV) of TS with time. No relationship to outcome was detected for TP or DPD. Expression of HER2 or p53 was a significant prognostic indicator in the tamoxifen alone group. TS, but not TP or DPD, may be a useful predictor of response to UFT therapy. After 2 years, the RPV of TS decreased with time, suggesting that 2 years of treatment with oral fluorouracil derivatives may be inadequate. Further studies are required to investigate this possibility.

Correspondence to: Dr Masakazu Toi, Department of Surgery (Breast Surgery), Graduate School of Medicine, Kyoto University, 54 Shogoin-Kawaramachi, Sakyo-ku, Kyoto 606-8507, Japan
E-mail: toi@kuhp.kyoto-u.ac.jp

Key words: tegafur-uracil, predictive factor, thymidylate synthase, thymidine phosphorylase, dihydropyrimidine dehydrogenase, breast cancer

Introduction

UFT is an oral formulation combining tegafur, a prodrug of 5-fluorouracil, with uracil, an inhibitor of dihydropyrimidine dehydrogenase (DPD), the rate-limiting enzyme governing the metabolism of 5-fluorouracil. Recently, many studies have demonstrated that adjuvant treatment with tegafur-uracil (UFT) is effective against lung cancer and other types of solid tumors (1-4). In breast cancer, the therapeutic usefulness of adjuvant chemotherapy with tegafur preparations has been studied in Japan and other countries for more than 20 years (5,6). Recently, Noguchi *et al* (7) reported the results of a pooled analysis of 6 randomized clinical trials in women with node-negative breast cancer. Their analysis demonstrated that survival was significantly longer in patients who received UFT than in those who did not. In addition, the effects of combined treatment with UFT and tamoxifen were found to be additive. These findings suggested that UFT may be useful for the management of primary breast cancer, although controlled studies with commonly used regimens for polychemotherapy, such as anthracycline plus cyclophosphamide (AC) and cyclophosphamide plus methotrexate plus fluorouracil (CMF), have yet to be reported.

Recent studies have shown that S-1, a combination of tegafur and 5-chloro-2,4-dihydropyrimidine (CDHP), a more potent inhibitor of DPD than uracil, has high antitumor activity against metastatic breast cancer (8). Other studies with 5-fluorouracil derivatives have demonstrated that combined treatment with capecitabine and docetaxel significantly prolongs survival among women with anthracycline-resistant breast cancer, as compared with docetaxel alone (9). Various

trials are now being performed in preoperative or postoperative settings (10). These drugs will most likely play an important role in the future treatment of breast cancer. The benefits of oral 5-fluorouracil derivatives would be further enhanced by the ability to predict response, thereby identifying patients most likely to benefit from treatment and increasing the benefit-risk ratio.

Various approaches have been proposed to predict the response to oral 5-fluorouracil derivatives. Experimental and clinical evidence has suggested that tumor levels of enzymes involved in nucleoside metabolism, such as thymidylate synthase (TS), thymidine phosphorylase (TP), and dihydropyrimidine dehydrogenase (DPD), may be useful for predicting the response to oral 5-fluorouracil derivatives. Predictive accuracy may be further enhanced by using these enzymes in conjunction with other molecular markers.

We retrospectively examined whether the expression of the 3 enzymes TS, DPD, and TP and that of the oncogene HER2 and the tumor-suppressor gene p53 in breast cancer tissue could be used to predict the response to treatment with tamoxifen plus UFT. Resected tissue specimens were obtained from women with breast cancer who were enrolled in the 3rd Adjuvant Chemo-Endocrine Therapy for Breast Cancer (ACETBC) trials, randomized controlled studies comparing tamoxifen alone with tamoxifen plus UFT after surgery.

Patients and methods

Combined analysis of three randomized trials. A meta-analysis of 5 randomized controlled trials (n=1987) performed by the ACETBC study group in Japan has shown that the reduction in the risk of recurrence after treatment with UFT was 21±11% (P=0.06) in women with stage I to IIIA breast cancer who underwent mastectomy (5).

Three of these trials examined the effect of adding UFT (300-400 mg/day) to tamoxifen (20-30 mg/day) in women with estrogen-receptor (ER)-positive tumors who postoperatively received adjuvant chemotherapy for 2 years. ER status was determined at each center. Either biochemical (enzyme immunoassay) or immunohistochemical techniques were used. In 2 of these trials, mitomycin C (10 mg/m^2) was given intravenously on the day of surgery. Combined analysis of these 3 trials (n=1225; median follow up, 5.7 years) was performed according to the method of Peto (Fig. 1). The reduction in the risk of recurrence after treatment with UFT plus tamoxifen was found to be 26±12% (p=0.037). Subset analyses of pooled data in the 3 trials showed that UFT was most effective in premenopausal women with metastases to the axillary lymph nodes (reduction in odds of recurrence, 35±17%).We retrospectively studied the predictive values of biomarkers of response in this patient subset.

Immunohistochemically studied biomarkers

Collection of tumor samples. A list of subjects was submitted to centers that had agreed to participate in this biomarker study and had registered at least 5 patients to the 3rd ACETBC study. All available paraffin-embedded samples were sent from the centers to the operational office by mail. The samples were stored at room temperature until predictive markers were evaluated.

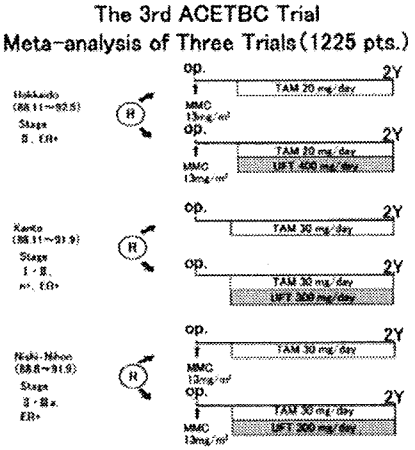

Figure 1. Protocols of the 3rd ACETBC trial.

Immunohistochemical labeling

Antibodies. TS polyclonal antibody RTSSA (dilution, 1:100; Taiho Pharmaceutical Co., Ltd., Tokyo, Japan), TP monoclonal antibody TMA-1 (dilution, 1:100; Taiho Pharmaceutical Co., Ltd.), DPD polyclonal antibody RDPDPA (dilution, 1:100; Taiho Pharmaceutical Co., Ltd.), HER2 polyclonal antibody A0485 (Dako, Carpinteria, CA, USA; dilution, 1:100), and p53 (DO7) monoclonal antibody (Novo-castra, Newcastle, UK; dilution, 1:40) were used for immunohistochemical analyses.

Immunohistochemical analyses. Immunohistochemical analyses were performed at a single central laboratory using the antibodies described above and mouse IgG (Dako) as negative control. An indirect avidin-biotin-peroxidase method was used. Briefly, deparaffinized tissue sections were treated with 0.3% hydrogen peroxide in methanol to block endogenous peroxidase activity. After washing with phosphate buffered-saline (PBS) containing 0.05% Tween-20, the sections were treated with 1.5% normal horse serum in PBS and incubated with each of the antibodies or with mouse IgG for 1 h at room temperature. The sections were washed again with PBS, incubated with biotinylated anti-mouse IgG (Dako) for 30 min, washed again with Tween-20-PBS, incubated with an elite ABC kit (Vector, Burlingame, CA, USA) for 30 min, and visualized with the use of 3,3'-diaminobenzidine tetrahydrochloride-hydrogen peroxide as chromogen. The sections were then counterstained with hematoxylin, dehydrated, and mounted.

Evaluation of staining. The slides were evaluated independently by 3 experienced pathologists (A.F., K.M., T.H.) blinded with regard to treatment group and outcome. Each pathologist evaluated TS, TP, and DPD on the basis of staining intensity of the cytoplasm, scored according to a 4-grade scale (0 to 3), and staining rate, also scored according to a 4-grade scale (≤25%, 0; >25% to ≤50%, 1; >50% to ≤75%, 2; and >75%, 3). The scores agreed on by 2 or more of the pathologists were adopted. Concordance rates of the evaluations among 2 or more pathologists were as follows: TS, staining intensity 95%, staining rate 80%; TP, staining intensity 92%, staining rate

Table I. Patients' characteristics in the biomarker study.

	TAM group (n=97)	UFT group (n=95)	p-value
Age			
≤50	89	89	0.78
>51	8	6	
Number of nodes involved			
1-3	65	73	0.15
≥4	32	22	
Tumor size			
<2 cm	23	24	0.87
≥2 cm	74	71	
TS expression			
Positive	57	48	0.31
Negative	40	47	
TP expression			
Positive	36	39	0.86
Negative	61	56	
DPD expression			
Positive	57	66	0.13
Negative	40	29	
HER2 expression			
Positive	14	14	1.00
Negative	83	81	
p53 expression			
Positive	30	33	0.85
Negative	67	62	

All patients had estrogen receptor-positive tumors and were pre-menopausal.

87%; and DPD, staining intensity 94%, staining rate 89%. The median score was adopted if all 3 pathologists disagreed on the score. Cases were considered positive if the staining intensity was ≥2, and the staining rate was 3 (staining rate, >75%).

HER2 was evaluated on the basis of staining of the membrane, and p53 was evaluated on the basis of staining of nuclei. The results were considered positive if the staining rate was ≥1%. The evaluation agreed on by 2 or more pathologists was adopted (concordance rates among the evaluations of the 3 pathologists were as follows: HER2, 89%; and p53, 72%).

Statistical analysis. Relapse-free survival was the outcome used to assess treatment efficacy and was defined as the interval elapsed between the date of surgery and the date of documented disease relapse or death. Relapse-free survival was calculated by the Kaplan-Meier method. Differences between groups in Kaplan-Meier estimates of relapse-free survival were evaluated with the log-rank test and generalized Wilcoxon test. Risk

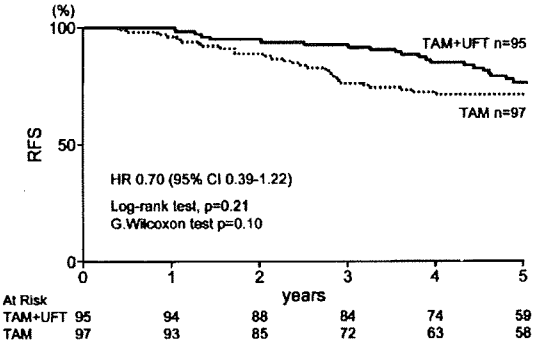

Figure 2. Relapse-free survival (RFS) according to study group (n=192).

ratios (RR) were estimated from Cox proportional-hazards regression models. No overall survival analysis was performed in the subgroups of patients identified by the evaluated biological markers because of the small numbers of events in each treatment group. Cox proportional-hazards regression models were also used to test for interactions between biomarkers and treatment.

Relative predictive values (RPV) were determined with use of the following equation, modified from the method described by Hayes (11): RPV for events in the tamoxifen + UFT group was compared with those in the tamoxifen alone group = Log (RR when tumors stained negatively for biomarkers/RR when tumors stained positively for biomarkers). Differences in distributions between groups were compared with the use of the χ^2 test. Differences were considered statistically significant when p-values were <0.05, and all reported p-values are two-tailed. All analyses were carried out with SAS software (version 6.12).

Results

Collection of samples. Samples collected from 192 (97 given tamoxifen and 95 given tamoxifen plus UFT) of the 204 women at the centers were assessable. There were no significant differences between the groups in demographic characteristics (age, tumor size, number of lymph node metastases) (Table I). The hazard ratio of the effect of adding UFT to tamoxifen was 0.70 (95% confidence interval, 0.39 to 1.22) (log-rank test, p=0.21; Wilcoxon test, p=0.10) (Fig. 2).

Expression of biomarkers. The rates of positive staining were as follows: TS, 55% (105/192); TP, 39% (75/192); DPD, 64% (123/192); HER2, 15% (28/192); and p53, 33% (63/192). The expression rates of these biomarkers were similar in the tamoxifen group and the tamoxifen plus UFT group (Table I).

Relation between relapse-free survival and expression of biomarkers in tumors. Demographic characteristics were similar in women whose tumors stained positively for each biomarker (TS, TP, or DPD) and those whose tumors stained negatively for each biomarker. Univariate analyses showed no significant differences in relapse-free survival between women whose tumors stained positively for TS, TP, or DPD and those whose tumors stained negatively for these 3

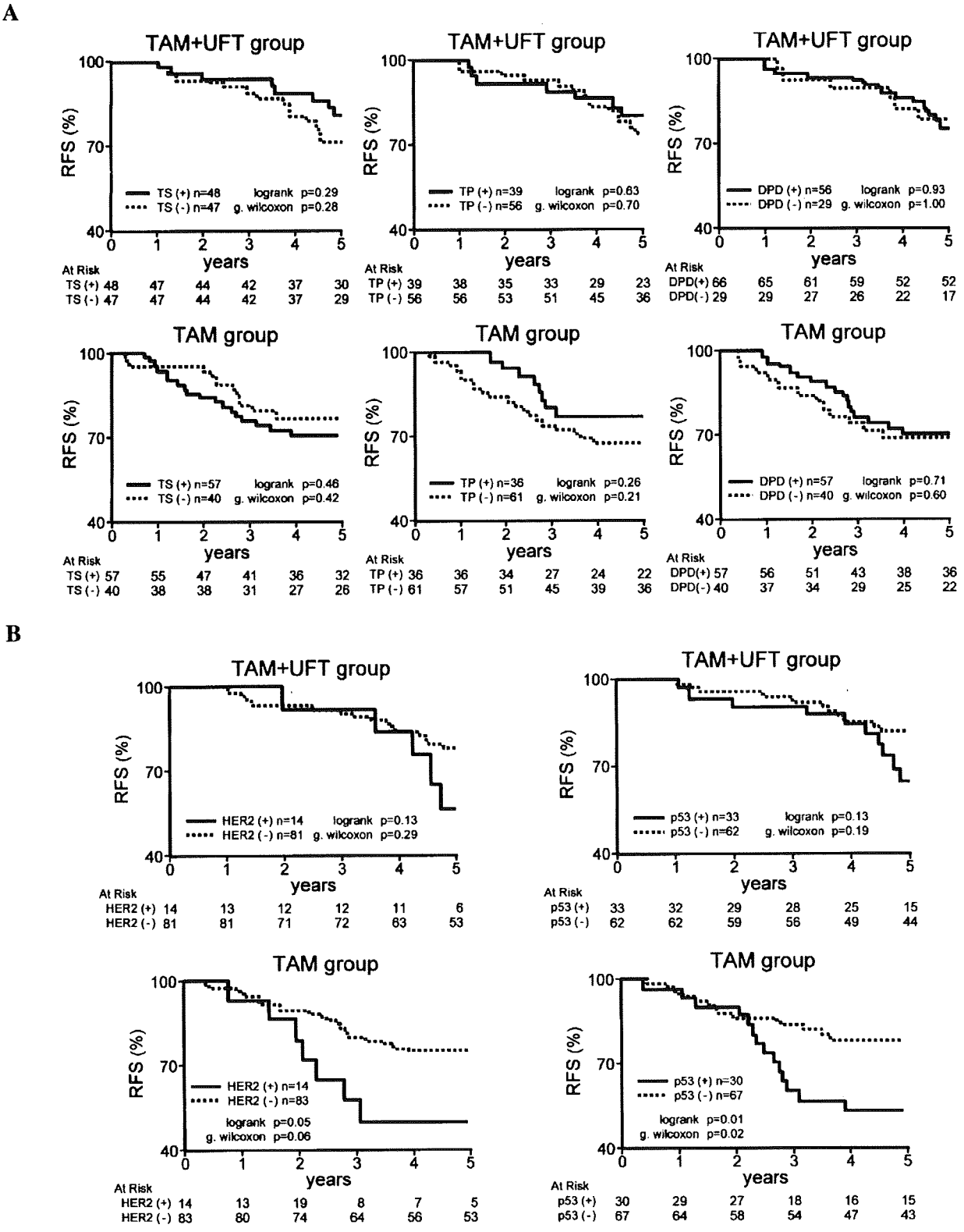

Figure 3. (A) Relation between relapse-free survival (RFS) and tumor expression of thymidylate synthase (TS), thymidine phosphorylase (TP), and dihydropyrimidine dehydrogenase (DPD) according to treatment. (B) Relation between relapse-free survival and tumor expression of HER2 and p53 according to treatment.

biomarkers in either treatment group. Women whose tumors stained positively for HER2 or p53 in the tamoxifen alone group had significantly poorer outcomes than those whose tumors stained negatively for these biomarkers. HER2 and p53 were not significant prognostic factors in the tamoxifen plus UFT group (Fig. 3).

Table II. Relative risk (TAM+UFT vs. TAM) according to biomarker expression.

Biomarker	Biomarker positive			Biomarker negative			Interaction
	RR	95% CI	p-value (G. Wilcoxon test)	RR	95% CI	p-value (G. Wilcoxon test)	p-value
TS	0.48	0.20-1.07	0.04	1.00	0.44-2.36	1.00	0.22
TP	0.80	0.28-2.23	0.60	0.66	0.33-1.30	0.124	0.76
DPD	0.75	0.37-1.52	0.29	0.61	0.21-1.56	0.222	0.73
HER2	0.59	0.17-1.86	0.19	0.72	0.37-1.37	0.220	0.77
p53	0.57	0.25-1.28	0.09	0.78	0.35-1.72	0.418	0.59

RR, relative risk by addition of UFT to TAM; TS, thymidylate synthase; TP, thymidine phosphorylase; DPD, dihydropyrimidine dehydrogenase.

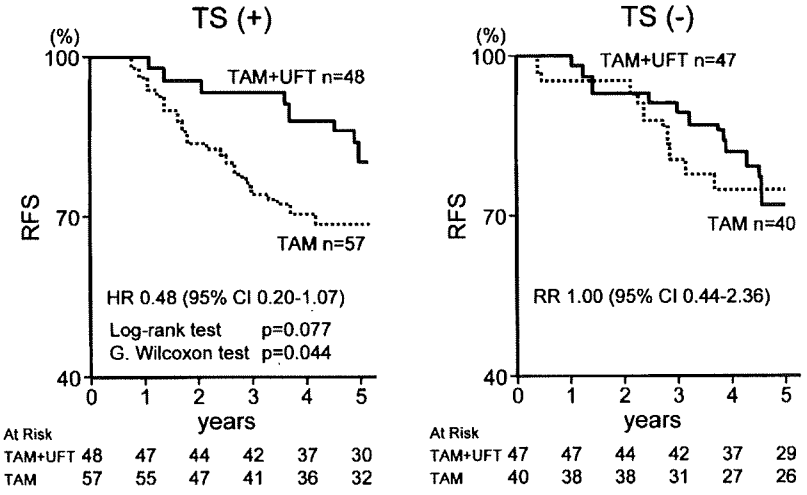

Figure 4. Comparison of relapse-free survival (RFS) between TAM and TAM+UFT treatment according to thymidylate synthase (TS) status.

Relation between expression of biomarkers in tumors and effect of adding UFT to tamoxifen

TS. In women with TS-positive tumors, the risk ratio of the effect of adding UFT to tamoxifen was 0.48 (95% confidence interval, 0.20 to 1.07), and response differed significantly between women given tamoxifen alone and those given tamoxifen plus UFT (p=0.04 by the generalized Wilcoxon test, p=0.08 by the log-rank test). In women with TS-negative tumors, however, there was no significant difference in response (hazard ratio, 1.00; 95% confidence interval, 0.44-2.36). Interaction testing showed that the expression of TS was not significantly related to the effect of UFT (p=0.22) (Fig. 3, Table II).

TP. The risk ratio of the effect of adding UFT to tamoxifen was 0.80 (95% confidence interval, 0.28-2.23) in women with TP-positive tumors and 0.66 (95% confidence interval, 0.33-1.30) in women with TP-negative tumors. There were no significant differences in response between the treatment groups. Interaction testing showed no significant relation between the expression of TP and the effect of UFT (p=0.76) (Table II).

DPD. The risk ratio of the effect of adding UFT to tamoxifen was 0.75 (95% confidence interval, 0.37-1.52) in women with DPD-positive tumors and 0.61 (95% confidence interval, 0.21-1.56) in those with DPD-negative tumors. There were no significant differences between the treatment groups. Interaction testing showed that the expression of DPD was not significantly related to the effect of UFT (p=0.73) (Table II).

HER2. The risk ratio of the effect of adding UFT to tamoxifen was 0.59 (95% confidence interval, 0.17-1.86) in women with HER2-positive tumors and 0.72 (95% confidence interval, 0.37-1.37) in those with HER2-negative tumors. There were no significant differences between the treatment groups. Interaction testing showed that the expression of HER2 was not significantly related to the effect of UFT (p=0.77) (Table II).

p53. The hazard ratio of the effect of adding UFT to tamoxifen was 0.57 (95% confidence interval, 0.25-1.28) in women with p53-positive tumors and 0.78 (95% confidence interval, 0.35-1.72) in women with p53-negative tumors. There were no significant differences between the treatment groups.