

HER2 in gastric cancer: Comparative analysis of three different antibodies using whole-tissue sections and tissue microarrays

Lucas Faria Abrahão-Machado, Alexandre Andrade dos Anjos Jácome, Durval Renato Wohnrath, José Sebastião dos Santos, Estela Cristina Carneseca, José Humberto Tavares Guerreiro Fregnani, Cristovam Scapulatempo-Neto

Lucas Faria Abrahão-Machado, Cristovam Scapulatempo-Neto, Department of Pathology, Barretos Cancer Hospital, Barretos, Sao Paulo 14784-400, Brazil

Alexandre Andrade dos Anjos Jácome, Department of Clinical Oncology, Mater Dei Hospital, Belo Horizonte, Minas Gerais 30140-092, Brazil

Durval Renato Wohnrath, Department of Gastrointestinal Surgical Oncology, Barretos Cancer Hospital, Barretos, Sao Paulo 14784-400, Brazil

José Sebastião dos Santos, Department of Surgery and Anatomy, University of São Paulo-Ribeirão Preto School of Medicine, Ribeirão Preto, Sao Paulo 14049-900, Brazil

Estela Cristina Carneseca, Biostatistician at Cancer Registry Department, Barretos Cancer Hospital, Barretos, Sao Paulo 14784-400, Brazil

José Humberto Tavares Guerreiro Fregnani, Researcher Support Center, Barretos Cancer Hospital, Barretos, Sao Paulo 14784-400, Brazil

Author contributions: Abrahão-Machado LF and Scapulatempo-Neto C designed the research; Abrahão-Machado LF, Scapulatempo-Neto C and Jácome AAA performed the research; Carneseca EC and Fregnani JHTG contributed with analytic tools; Abrahão-Machado LF and Scapulatempo-Neto C analyzed the data; Abrahão-Machado LF wrote the paper; Wohnrath DR and Santos JS revised the final version.

Correspondence to: Lucas Faria Abrahão-Machado, MD, Department of Pathology, Barretos Cancer Hospital, Rua Antenor Duarte Vilela, 1331, Barretos, Sao Paulo 14784-400, Brazil. lucasfariamachado@yahoo.com

Telephone: +55-17-3321660-6987 Fax: +55-17-3321660-7117

Received: June 5, 2013

Revised: July 12, 2013

Accepted: July 17, 2013

Published online: October 14, 2013

Abstract

AIM: To compare the performance of three commercially available anti-human epidermal growth factor re-

ceptor 2 (HER2) antibodies in whole-tissue sections and tissue microarrays (TMAs) of a series of gastric tumors.

METHODS: We present a comparative analysis of three anti-HER2 antibodies (HercepTest, 4B5 and SP3) using TMA and whole-tissue sections prepared from the same paraffin blocks of 199 gastric adenocarcinomas operated upon between January 2004 and December 2008 at a Brazilian cancer hospital. The data on the patients' age, sex, the anatomical location of the tumor and the Lauren's histological classification were collected from clinical and pathological records. The immunohistochemical (IHC) results were examined by two pathologists and the cases were classified as positive (3+), equivocal (2+) and negative (0 or 1+), according to the criteria of the IHC scoring system of gastric cancer. TMAs and whole-tissue sections were evaluated separately and independently. All cases yielding discordant IHC results and/or scored as 2+ were subjected to dual-color *in situ* hybridization in order to determine the final HER2 status. Besides determining the sensitivity and predictive value for HER2-positive status, we measured the accuracy of each antibody by calculating the area under the receiver operating characteristic (ROC) curve. The agreement between the results obtained using the TMAs and those obtained using the whole-tissue sections was assessed by means of Kappa coefficient.

RESULTS: Intratumoral heterogeneity of HER2 expression was observed with all antibodies. HER2-positive expression (3+) in the whole-tissue sections was observed in 23 cases (11.6%) using the 4B5 antibody, in 18 cases (9.1%) using the SP3 antibody and in 10 cases (5.1%) using the HercepTest antibody. In the TMAs, 11 positive cases (5.6%) were identified using SP3 antibody, 9 (4.6%) using the 4B5 antibody and 6 (3%) using the

HercepTest antibody. The sensitivity using whole-tissue sections and TMA, respectively, was 95.2% and 42.9% with 4B5, 90.5% and 66.7% with SP3 and 47.6% and 42.9% with HercepTest. The accuracy, calculated from the area under the ROC curve, using whole-tissue sections and TMA, respectively, was 0.91 and 0.79 by 4B5, 0.86 and 0.80 by SP3 and 0.73 and 0.71 by HercepTest. The concordance of the results obtained using whole-tissue sections and TMA was 97.4% (Kappa 0.75) using HercepTest, 85.6% (Kappa 0.56) using SP3 and 84.1% (Kappa 0.38) using 4B5.

CONCLUSION: The use of the 4B5 antibody on whole-tissue sections was the most accurate IHC method for evaluating HER2 expression in gastric adenocarcinoma.

© 2013 Baishideng. All rights reserved.

Key words: Gastric cancer; Human epidermal growth factor receptor 2; Immunohistochemistry; Whole-tissue sections; Tissue microarray; Trastuzumab

Core tip: This is the first study to compare the three widely used anti-human epidermal growth factor receptor 2 (HER2) antibodies 4B5, SP3 and HercepTest in tissue microarrays and whole-tissue sections prepared from paraffin blocks of a single series of gastric tumors. We aimed to find the best method to assess HER2 expression in gastric cancer, facilitating the choice of the antibody with the greatest ability to identify the most patients who could benefit from the use of trastuzumab. Besides, we demonstrated that HER2 expression in small samples of gastric cancer (such as tissue microarrays and biopsies) should be evaluated cautiously because these tumors exhibit intratumoral heterogeneity that may influence the results.

Abrahão-Machado LF, Jácome AAA, Wohnrath DR, Santos JS, Carneseca EC, Fregnani JHTG, Scapulatempo-Neto C. HER2 in gastric cancer: Comparative analysis of three different antibodies using whole-tissue sections and tissue microarrays. *World J Gastroenterol* 2013; 19(38): 6438-6446 Available from: URL: <http://www.wjgnet.com/1007-9327/full/v19/i38/6438.htm> DOI: <http://dx.doi.org/10.3748/wjg.v19.i38.6438>

INTRODUCTION

The incidence of gastric cancer (GC) is gradually decreasing; however, it remains one of the leading causes of cancer-related death worldwide because the vast majority of GC patients are diagnosed with advanced disease^[1-4]. Despite the improvement in surgical techniques and the use of multimodal treatments, the prognosis for GC is generally poor and treatment continues to be a challenge for physicians^[1,4]. Recently, several oncogenes and tumor suppressor genes were studied in an attempt to clarify the process of gastric carcinogenesis, and specific monoclonal antibodies were developed as a potential form of adjuvant treatment for patients with advanced disease.

The HER2 (CerbB-2) or human epidermal growth factor receptor 2 (HER2) gene is a proto-oncogene located on chromosome 17q21 that encodes a transmembrane protein that is a member of the HER receptor family. These receptors possess tyrosine kinase activity and are typically involved in signal transduction pathways that lead to cell growth and differentiation^[5]. Amplification of the *HER2* gene and overexpression of its product have been identified in several tumors and have been widely studied in breast cancer^[6]. In GC, however, the reported frequency of HER2 overexpression ranges from 8.2% to 53.4%, and its clinical significance and prognostic value remain controversial, although HER2-positive tumors are usually associated with more aggressive biological behavior and tumor recurrence^[7-13]. A recent meta-analysis showed that in 7 of the 15 papers evaluated, HER2 positivity was correlated with a worse prognosis^[14].

New advances in molecular targeting therapy have identified HER2 as an important target for anti-cancer therapy of gastric tumors. The ToGa study recently indicated improved survival of patients with advanced GC who were treated with trastuzumab (a chimeric anti-HER2 targeted drug) combined with chemotherapy compared with those treated with chemotherapy alone^[15]. This randomized clinical trial achieved the longest median survival to date of patients with advanced gastric carcinomas. The mechanism by which trastuzumab acts is not completely understood, but the likely possibilities are that it prevents the dimerization of HER2 with other members of the HER family, activates the immune response by promoting antibody-dependent cell-mediated toxicity and induces endocytosis of HER2^[16,17]. Given the demonstration of its clinical benefits and its approval for use in systemic therapy by the Food and Drug Administration (FDA), trastuzumab is the new standard treatment option for patients with HER2-positive advanced GC. Therefore, it is crucial to determine the HER2 status of GCs to select patients who may benefit from this promising targeted therapy.

Several assays are available to determine HER2 status; however, many of them require fresh tissue, involve complicated procedures and are costly. The most commonly used method is immunohistochemistry (IHC), which is a low-cost technique that can be performed on small samples, even formalin-fixed and paraffin-embedded tissues. Fluorescent *in situ* hybridization (FISH) is considered the gold standard and can be used to analyze this type of sample. However, because of its higher cost and the need for a fluorescence microscope, as well as the high concordance between FISH and IHC reported in literature^[18-21], generally only equivocal cases are subjected to FISH. An alternative for equivocal cases is provided by the use of other *in situ* hybridization methods such as silver *in situ* hybridization (SISH), including dual-color *in situ* hybridization (DISH), which allows the use of an ordinary light microscope and has shown excellent correlation with results obtained using FISH^[18,22,23].

Although a widely used and FDA/CAP-approved IHC scoring system already exists for HER2 in breast

cancer, it was necessary to develop a suitable scoring system for gastric tumors, mainly because of morphological differences and the intratumoral heterogeneity of HER2 expression in GC^[8,9,11,18,24]. The system proposed by Hoffmann *et al*^[9] for GC and incorporated as standard by CAP and FDA differentiates between surgical specimens and biopsies.

Currently, commercially available IHC antibodies include the HercepTest and A0485 (Dako, Glostrup, Denmark) rabbit polyclonal antibodies, the SP3 (Labvision; Thermo Fisher Scientific, Fremont, CA, United States) and 4B5 (Ventana Medical Systems, Tucson, AZ, United States) rabbit monoclonal antibodies and the CB11 mouse monoclonal antibody (Novocastra, Newcastle upon Tyne, England). Only the HercepTest, 4B5 and CB11 antibodies are approved by the FDA, although the international literature also shows high-quality of the SP3 antibody in samples of breast cancer^[20,25].

In the present study, HER2 expression in 199 GC was investigated by IHC on whole-tissue sections and tissue microarrays (TMA) using HercepTest, 4B5 and SP3. To date, no published results have compared these three antibodies. Moreover, this is the first study of GC to compare HER2 expression using both TMAs and whole-tissue sections prepared from samples of the same paraffin blocks, *i.e.*, the same tumors. All cases yielding divergent IHC results or results considered equivocal (2+) were subjected to DISH. We hypothesized that if the TMA samples were considered to be biopsies because they are small tissue samples, the reproducibility of the HER2 scoring system for GC could be tested using the two types of specimens.

Given that HER2 expression in the stomach is heterogeneous, the main purpose of our study was to compare the performance of three commercially available anti-HER2 antibodies. Furthermore, we aimed to determine the concordance of results obtained from whole-tissue sections and TMAs from the same tumors to evaluate the feasibility of TMA as an alternative method for assessing HER2 expression in GC.

MATERIALS AND METHODS

Patients

In the present study, we selected 199 cases of surgically resected primary gastric or gastro-esophageal adenocarcinomas. All the patients were operated upon between January 2004 and December 2008 at the Barretos Cancer Hospital. Clinical data were collected from medical charts and pathology reports, including sex and patient age as well as the anatomical location of the tumor and its Laurén histological classification.

TMA construction and IHC

Paraffin blocks containing representative samples of the tumors were selected by reviewing all of the hematoxylin and eosin (HE) stained slides. For the TMAs, two tissue cores with a diameter of 0.6 mm were extracted from

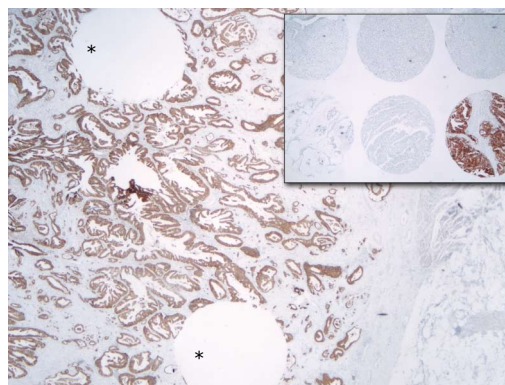


Figure 1 Photomicrograph of immunohistochemistry ($\times 100$) in a whole-tissue section. Asterisks indicate two round voids where the tissue microarray cores (inset) were extracted.

each tumor using the TMA arrayer MTA1 (Estigen, Tartu, Estonia). The tumor cores were sequentially placed in molds, embedded in paraffin and cooled to form the tissue array blocks. Each TMA also contained various non-gastric tissue samples as control tissues.

Sections of a thickness of 4 μm were obtained from the whole-tissue paraffin blocks and TMA blocks and used for IHC (Figure 1). The slides were stained using automatic staining devices: the Benchmark XT (Ventana Medical Systems, Tucson, AZ) for the 4B5 and SP3 antibodies and the Autostainer Link 48 (Dako, Glostrup, Denmark) for the HercepTest. After antigen retrieval processing for 60 min (at pH 8.4), 4B5 (prediluted form as provided by the manufacturer) and SP3 (diluted 1:100) were applied for a 32 min incubation period. Antibody visualization was enabled using the Ventana Ultraview DAB detection kit. The HercepTest was performed according to the manufacturer's guide provided with the kit, using the prediluted "ready-to-use" form for all of the steps and incubation periods preprogrammed in the stainer software. The kit also contained the visualization reagent. All the slides were subsequently counterstained with hematoxylin.

The criteria suggested by Hoffmann *et al*^[9] were used to evaluate the expression of HER2. Sections of the surgical specimens were considered HER2-positive (3+) when strong complete or basolateral membranous staining was detected in $\geq 10\%$ of the neoplastic cells; equivocal (2+) when moderate/weak complete or basolateral membranous staining was detected in $\geq 10\%$ of the cells; 1+ (negative) when the staining was weak or detected in only one part of the membrane in $\geq 10\%$ of the cells and 0 (also negative) in cases in which there was no membranous staining or staining of $< 10\%$ of the tumor cells. The criteria for evaluating biopsies were applied to the TMAs, and the percentage above (10%) was replaced by a cellular group of at least 5 cells. Full-tissue sections or TMA cores with excessive tissue fragmentation, scant invasive tumor and excessive cytoplasmic or background staining were rejected, and IHC was repeated on more suitable samples.

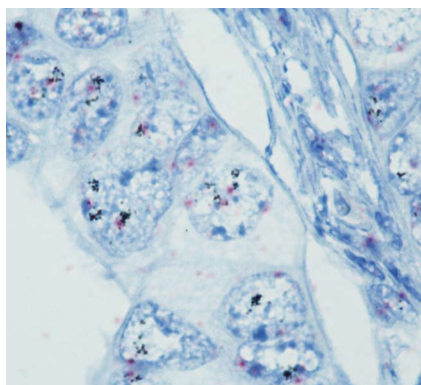


Figure 2 Dual-color *in situ* hybridization ($\times 1000$), human epidermal growth factor receptor 2 amplification. Black dots: Human epidermal growth factor receptor 2 gene; Pink dots: Chromosome 17.

***In situ* hybridization**

DISH was performed in all cases that were scored 2+ in either of the samples stained with any of the antibodies, in accordance with the guidelines recommended by CAP and routine laboratory practice. In addition, all of the cases with discordant IHC results were tested by DISH. The tissue sections used for DISH were obtained from the whole-tissue paraffin blocks. HER2 DISH was performed using the Ventana Benchmark XT-machine (Ventana Medical System, Tucson, AZ) following its standardized protocol. DISH is dual-color *in situ* hybridization in which the *HER2* gene is labeled using silver to produce a black dot, and the centromere of chromosome 17 (Chr 17) is labeled with alkaline phosphatase to produce a pink dot. Therefore, the *HER2* gene and Chr 17 are both simultaneously stained on the same slide (Figure 2). The DISH slides were examined using an HE slide to assist with the tumor location and morphology within each section. At least 40 tumor cell nuclei were scored for the Chr 17 signal and HER2 signal in different areas of the tumor. Only nuclei displaying both signals were scored, and a HER2/Chr 17 ratio was obtained for each specimen. HER2 amplification was defined as a ratio of HER2/Chr 17 ≥ 2 . Chromosome 17 polysomy was defined as ≥ 3 Chr 17 signals per cell on average^[26].

Statistical analysis

The statistical analysis was conducted using SPSS version 19.0 software (SPSS Inc., Chicago, IL). The IHC results were compared with a final variable of positivity or negativity for HER2 protein expression. In equivocal cases or cases of nonconcordant results obtained with the three antibodies, *HER2* gene amplification was assayed by DISH, and the DISH results determined the definitive HER2 status. Therefore, the results considered final, *i.e.*, the gold standard for statistical analysis, were those that were identical for the three antibodies, in addition to the results of DISH. The area under the receiver operating characteristic (ROC) curve (AUC) for each test was used to measure the accuracy of antibody labeling. To verify the agreement between the results obtained using the TMAs and those obtained using the whole-tissue sec-

tions, we employed the Kappa coefficient^[27]. We defined $P < 0.05$ as statistically significant.

RESULTS

Patients and tumor characteristics

There were 123 male and 76 female patients, and their age ranged from 27 to 87 years (median: 60.7 years). The location of the tumor was in the antrum in 86 cases (43.2%), in the body in 26 cases (13%), in the fundus in 2 cases (1%), in the cardia in 38 cases (19%) and multicentric (in more than 2 regions) in 47 cases (23.6%). Fifty-five cases were histologically classified as the diffuse type (27.6%), 123 as the intestinal type (61.8%), 17 as the mixed type (8.5%) and four as not otherwise classified (2%).

Immunohistochemistry

HER2-positive expression (3+) in the whole-tissue sections was observed in 10 cases (5.1%) using the HerceptTest, in 23 cases (11.6%) using the 4B5 antibody and in 18 cases (9.1%) using the SP3 antibody. Using the TMAs, 6 HER2-positive cases (3%) were identified using the HerceptTest, 9 (4.6%) using the 4B5 antibody and 11 (5.6%) using SP3 antibody. The immunohistochemistry results are shown in Table 1.

The HerceptTest demonstrated the lowest number of positive cases in both the whole-tissue sections and the TMAs. The SP3 antibody yielded the highest number of equivocal (2+) cases for both types of samples. The frequency of a score of 2+ was higher among the whole-tissue sections than among TMAs, except when using the HerceptTest, which showed the opposite pattern.

The overall concordance between the results obtained using the TMAs and those obtained using the whole-tissue sections was 97.4% with the HerceptTest, 84.1% with the 4B5 antibody and 85.6% with the SP3 antibody. According to the values of the Kappa coefficient, HerceptTest provided substantial agreement between the TMAs and whole-tissue sections, the SP3 antibody provided moderate agreement and the 4B5 antibody provided fair agreement (Table 2).

Stronger membrane staining in positive cases was observed for the 4B5 antibody than for the other two antibodies (Figure 3). The diffuse cytoplasmic staining in the gastric foveolar epithelium and intestinal metaplasia that was observed when using 4B5 antibody was less pronounced when using the SP3 antibody and not observed when using the HerceptTest antibody. Heterogeneous HER2 expression within the tumors was observed with all antibodies (Figure 4). All the positive cases were classified as intestinal type. Nuclear staining with the 4B5 and SP3 antibodies was observed in some of the diffuse adenocarcinomas.

DISH and final HER2 status

Cases with divergent results and those considered equivocal by IHC (scored as 2+) were subjected to DISH; there

Table 1 Results of HER2 immunostaining using the three antibodies on whole-tissue sections and tissue microarrays *n* (%)

Score	Whole-tissue sections			TMAs		
	HercepTest	4B5	SP3	HercepTest	4B5	SP3
0	179 (90.9)	125 (63.1)	128 (65.2)	185 (93.5)	174 (88.8)	162 (82.6)
1+	7 (3.5)	30 (15.2)	17 (8.5)	3 (1.5)	10 (5.1)	8 (4.1)
2+	1 (0.5)	20 (10.1)	34 (17.2)	4 (2.0)	3 (1.5)	15 (7.7)
3+	10 (5.1)	23 (11.6)	18 (9.1)	6 (3.0)	9 (4.6)	11 (5.6)
Total	197 (100.0)	198 (100.0)	197 (100.0)	198 (100.0)	196 (100.0)	196 (100.0)

TMA: Tissue microarray.

Table 2 Concordance between the tissue microarrays and whole-tissue sections staining results using the HercepTest, 4B5 and SP3 antibodies

Antibody	Overall concordance	Kappa coefficient (95%CI)
HercepTest	97.40%	0.75 (0.54-0.96)
4B5	84.10%	0.38 (0.22-0.53)
SP3	85.60%	0.56 (0.43-0.70)

were 58 (29.1%) such cases, of which 14 cases (24.1%) exhibited *HER2* gene amplification and 44 cases (76.9%) did not exhibit *HER2* gene amplification. Chr 17 polysomy was present in 5 cases (8.6%), but it was not related to amplification in our study. Table 3 shows the *HER2* gene status in the cases that presented an IHC score of 2+.

A final positive *HER2* status (either by IHC or DISH) was obtained in 20 of the 199 cases tested (10%). The sensitivity using whole-tissue sections was 47.6% with HercepTest, 95.2% with 4B5 and 90.5% with SP3 in cases with an immunoscore of 2+/3+. The sensitivity using TMA was 42.9% with HercepTest, 57.1% with 4B5 and 66.7% with SP3 (Table 4).

Table 5 demonstrates the accuracy of each antibody. The 4B5 and SP3 antibodies gave similar AUC values in whole-tissue sections (Table 5), and both were significantly more accurate than HercepTest ($P = 0.002$ and 0.035 , respectively). Although the 4B5 and SP3 antibodies both gave greater AUC values than HercepTest in TMAs, the difference was not statistically significant. Based on the AUC of each antibody for both types of samples and the respective P values (Table 6), we determined that the use of the 4B5 antibody on whole-tissue sections was the most accurate method. SP3 staining of whole-tissue sections was also more accurate than the HercepTest using TMAs ($P = 0.013$).

Histological type and anatomical location

Of the 123 adenocarcinomas of the intestinal type, nine (7.3%) were given a final positive *HER2* status. None of the 76 cases of the other histological types was positive.

Of the 86 carcinomas located in the antrum, nine (10.4%) had a final positive *HER2* status. Eight (21%) of the 38 tumors situated in the cardia and three (6.4%) of the 47 multicentric tumors were positive.

DISCUSSION

With the demonstration of the benefits of trastuzumab therapy for advanced GC^[15], the clinical demand for *HER2* assessment is rapidly increasing. The use of trastuzumab in association with platinum and capecitabine or 5-FU for *HER2*-positive GC has shown the longest median survival in GC patients^[15]. Because IHC appears to be the easiest, least expensive and most widely used method, our goal was to compare three commercially available antibodies. Although the use of different clones can be problematic for GC, only two other reports in the literature have compared the performance of *HER2* antibodies^[18,28]. An ideal antibody test would be sufficiently sensitive to identify all possible treatment candidates and would have a low false-positive rate to minimize over-treatment.

Variability in performance among commercially available anti-*HER2* antibodies has been demonstrated in several studies, although most of these studies were performed in breast tumors^[29,30]. Cho *et al.*^[28] compared the HercepTest, A0485, 4B5 and CB11 antibodies in TMAs of gastric carcinomas, and they found that the sensitivity and specificity were 78.9% and 96% with HercepTest, respectively, 86.5% and 94.4% with the A0485 antibody, 76.3% and 95.6% with 4B5 and 60.5% and 98.4% with the CB11 antibody. Boers *et al.*^[18] tested the SP3 and 4B5 antibodies on biopsy specimens of gastro-esophageal adenocarcinomas, and they showed sensitivities of 77% and 96% as well as specificities of 100% and 98.4%, respectively. The latter result is consistent with the results we obtained when comparing the SP3 and 4B5 antibodies. In the present study, however, the difference in sensitivity among antibodies was much higher, as the sensitivity ranged from 47.6% to 95.2% when whole-tissue sections were analyzed and ranged from 42.9% to 66.7% when TMAs were analyzed. The sensitivity of HercepTest was far lower than that of the other two tests, which contributed to this difference. The specificity ranged from 81.2% to 99.4% for the whole-tissue sections and from 93.1% to 100% for the TMAs. For all the antibodies, lack of staining (score 0) was highly predictive of a negative/nonamplified case as confirmed by DISH, and positive staining (score 3+) was highly predictive of an amplified case as confirmed by DISH.

The 4B5 and SP3 antibodies exhibited similar perfor-

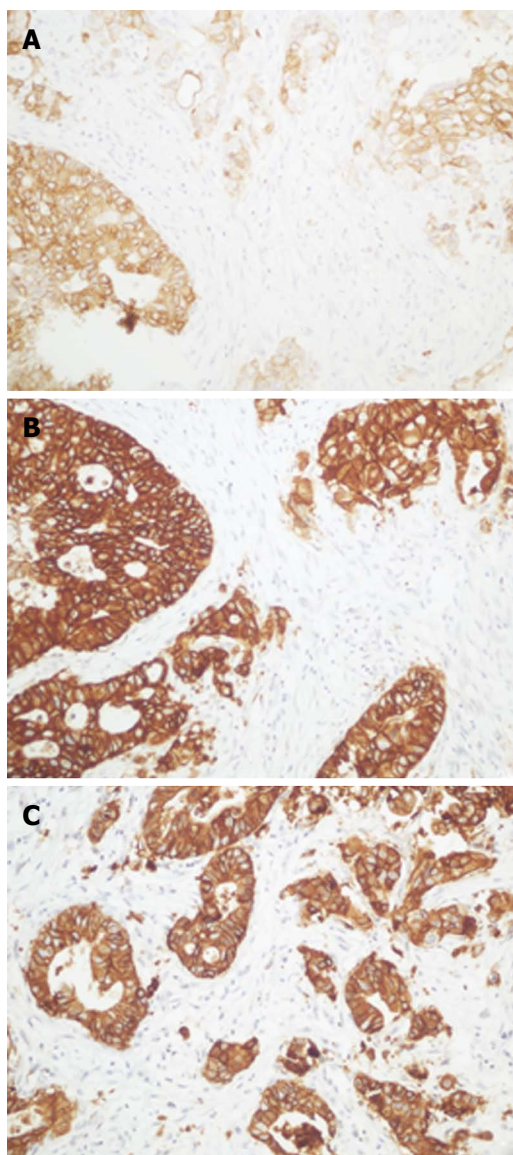


Figure 3 Comparison of positivity (3+) using the HercepTest (A), 4B5 (B) and SP3 (C) antibodies ($\times 200$).

mance, with high NPV values and AUC values that indicated higher accuracy, compared to HercepTest, although the difference was only statistically significant ($P < 0.05$) for whole-tissue sections. Even though HercepTest had high values for PPV and specificity, it presented the lowest sensitivity. Thus, this antibody provided the highest number of tumors with immunoscores of 0 or 1+ that were positive for HER2 amplification using DISH, which agrees with the report of Dekker *et al.*^[29] for breast tumors. Thus, in our view, HercepTest is not the best antibody to use as a first-line test to assess the HER2 status of GC because the ideal antibody should be highly sensitive, even though high sensitivity could increase the number of equivocal cases (2+) and the need for *in situ* hybridization tests. The 4B5 and SP3 antibodies were highly sensitive; therefore, these two antibodies appear to be more reasonable for first-line tests than HercepTest.

Another issue that we wanted to address was the use

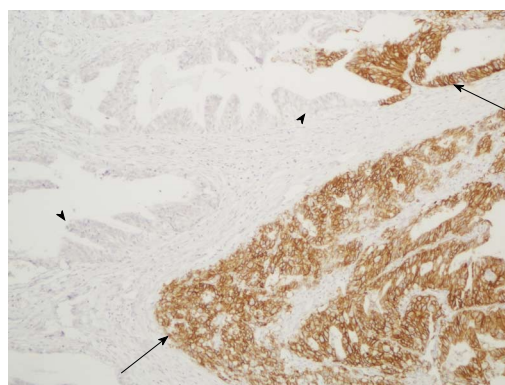


Figure 4 Representative image of the intratumoral heterogeneity of HER2 expression ($\times 100$). Arrows indicate areas with strong continuous membranous staining (score 3+) and arrowheads indicate negative areas (score 0).

of TMAs for assessment of HER2 and, by analogy, the reliability of testing endoscopic biopsies. The use of TMAs permits the inclusion of several different tumors in the same assay on a single slide. This cost-effective technique has become a standard procedure for many contemporary IHC studies. Dekker *et al.*^[29] found that TMAs were reliable for retesting large volumes of previously HER2-classified breast carcinomas. Similarly, Drev *et al.*^[31] observed a high concordance between whole-tissue sections and TMAs for breast tumors. Despite these favorable results for TMAs of breast cancers, HER2 assessment of gastric adenocarcinomas is more problematic. The obvious disadvantage of TMAs is that this preparation enables the analysis of only a limited sample of the tumor and for GCs, TMAs are even more unfavorable because of the generally observed heterogeneous expression of HER2 within these tumors.

Intratumoral heterogeneity can be defined as areas with different HER2 scores within the same tumor. This is the predominant pattern for GCs but not for breast tumors, and may thus cause sampling error when randomly sampled TMA cores of GCs are used^[8,9,11,24]. The difference in HER2 status between primary and metastatic tumor samples is still a matter of debate, and although the few studies in the scientific literature have demonstrated high concordance among the results obtained using these samples, some cases gave discordant results^[32,33], which suggests an effect of intratumoral heterogeneity.

Conspicuously heterogeneous HER2 staining on the whole-tissue sections and TMA core samples from most of the tumors was noted in our study. To minimize the discrepancy with results obtained with whole-tissue sections, we included two core samples from different areas of each tumor in the TMAs. Regardless, the TMA staining was much less sensitive than the staining of whole-tissue sections (mean values for the antibodies: 55.5% *vs* 77.7%, respectively). Although the HercepTest provided greater agreement between the TMAs and the whole-tissue sections, with a substantial Kappa value, its sensitivity was low for both types of sample, as mentioned above. The difference in HER2 expression detected in

Table 3 Human epidermalgrowth factor receptor 2 gene status assessed by dual-color *in situ* hybridization in cases with an immunohistochemistry score of 2+ with HercepTest, 4B5 and SP3 staining on whole-tissue sections and tissue microarrays *n* (%)

HER2 gene status	Whole-tissue sections			TMAs		
	HercepTest	4B5	SP3	HercepTest	4B5	SP3
Amplified	0 (0)	6 (30)	6 (17.6)	3 (75)	3 (100)	4 (26.6)
Not amplified	1 (100)	14 (70)	28 (82.4)	1 (25)	0 (0)	11 (73.4)
Total	1 (100)	20 (100)	34 (100)	4 (100)	3 (100)	15 (100)

TMA: Tissue microarray; HER2: Human epidermalgrowth factor receptor 2.

Table 4 Specificity, sensitivity, positive and negative predictive values and the area under the receiver operating characteristic curve of each antibody according to the final human epidermalgrowth factor receptor 2 status

		HercepTest (95%CI)	4B5 (95%CI)	SP3 (95%CI)
TMA	Sensitivity	42.9 (21.7-64.0)	57.1 (35.9-78.3)	66.7 (46.5-86.8)
	Specificity	99.4 (98.3-100.0)	100.0	93.1 (89.4-96.9)
	PPV	90.0 (71.4-100.0)	100.0	53.8 (34.7-73.0)
	NPV	93.6 (90.1-97.1)	95.1 (92.0-98.2)	95.9 (92.9-98.9)
	AUC	0.71 (0.60-0.782)	0.79 (0.68-0.89)	0.80 (0.69-0.90)
Whole-tissue sections	Sensitivity	47.6 (26.2-68.9)	95.2 (86.1-100.0)	90.5 (77.9-100.0)
	Specificity	99.4 (98.3-100.0)	87.0 (82.0-91.9)	81.2 (75.5-87.0)
	PPV	90.9 (73.9-100.0)	46.5 (31.6-61.4)	36.5 (23.4-49.6)
	NPV	94.1 (90.7-97.5)	99.3 (98.1-100.0)	98.6 (96.7-100.0)
	AUC	0.73 (0.63-0.84)	0.91 (0.86-0.96)	0.86 (0.78-0.93)

PPV: Positive predictive value; NPV: Negative predictive value; AUC: Area under the receiver operating characteristic curve; TMA: Tissue microarray.

Table 5 Accuracy of the three antibodies in the whole-tissue sections and the tissue microarrays

		HercepTest	4B5	SP3
Whole-tissue sections	AUC	0.73	0.91	0.78
	HercepTest	-	<i>P</i> = 0.002	<i>P</i> = 0.035
	4B5	<i>P</i> = 0.002	-	<i>P</i> = 0.265
TMAs	AUC	0.71	0.79	0.8
	HercepTest	-	<i>P</i> = 0.058	<i>P</i> = 0.075
	4B5	<i>P</i> = 0.058	-	<i>P</i> = 0.714
	SP3	<i>P</i> = 0.075	<i>P</i> = 0.714	-

AUC: Area under the receiver operating characteristic curve; TMA: Tissue microarray.

our study between TMAs and whole-tissue sections was caused by the prominent heterogeneity of HER2 staining. The 4B5 antibody results in whole-tissue sections were significantly different from the results of the three antibodies in TMAs. Because 4B5 antibody staining of whole-tissue sections had the highest accuracy, which was much different from that obtained for TMAs, our results suggest that TMA staining is less accurate and lacks sufficient sensitivity to reliably assess the HER2 status in GC. Two tissue cores for the TMA were definitely not sufficient to prevent sampling error and minimize false results because of the intratumoral heterogeneity and the small amount of tissue in the cores. Therefore, studies in the literature that used TMA to test the HER2 status of GCs must be carefully analyzed, and it must be noted that this technique does not seem to reflect the real status of the HER2 gene in GCs.

Table 6 Comparison of the areas under the receiver operating characteristic curve according to the type of sample (whole-tissue section and tissue microarray) and the antibodies

		Whole-tissue sections			
		HercepTest	4B5	SP3	
TMAs	AUC	0.73	0.91	0.78	
	HercepTest	0.71	<i>P</i> = 0.572	<i>P</i> = 0.001	<i>P</i> = 0.013
	4B5	0.79	<i>P</i> = 0.289	<i>P</i> = 0.027	<i>P</i> = 0.201
	SP3	0.80	<i>P</i> = 0.265	<i>P</i> = 0.034	<i>P</i> = 0.244

AUC: Area under the receiver operating characteristic curve; TMA: Tissue microarray.

Endoscopic biopsies with few fragments, such as those used in the TMA, may underestimate the incidence of HER2-amplification, as Yang *et al*^[24] demonstrated in a recent study in which large surgical specimens had higher rates of HER2 positivity than biopsy specimens. We believe that endoscopic biopsies are not optimal to identify the maximal number of patients who could be eligible for treatment. Therefore, to represent the tumor better and reduce misinterpretation, it is important to examine as many pieces of a biopsy as possible. We also suggest that all excisional specimens that had a previous HER2-negative result in a biopsy specimen should be retested to increase the chance of classifying the tumor as HER2-positive. Because the intratumoral heterogeneity of HER2 expression also seems to cause divergent results for primary and metastatic tumor samples^[32], is highly advisable to analyze the HER2 status of both primary and metastatic specimens when possible.

Thus, among the HercepTest, 4B5 and SP3 antibodies, HercepTest was the least sensitive and therefore had the lowest ability to identify a large number of patients eligible for trastuzumab treatment. According to our results, the most accurate IHC method to assess HER2 expression in GC is the use of the 4B5 antibody on whole-tissue sections. Intratumoral heterogeneity appears to be a major limitation for the use of TMA because TMA does not reflect the true HER2 status of many tumors. Because the number of cells that respond to a targeted therapy directly affects the tumor's responsiveness to treatment, there must be a great difference between cases that are diffusely positive and those that are only focally positive but still meet the criteria of positivity. Given the promising results from the use of trastuzumab and the particularities of HER2 expression in the stomach, further trials are needed to determine the clinical significance of the intratumoral heterogeneity and its impact on treatment outcome.

COMMENTS

Background

The vast majority of patients with gastric cancer are diagnosed with advanced disease and the prognosis is generally poor. Human epidermal growth factor receptor 2 (HER2)-positive tumors are usually associated with more aggressive biological behavior and recurrence. In view of the recently demonstrated clinical benefit of the anti-HER2 drug trastuzumab in the treatment of advanced gastric cancer, reliable HER2 testing is of key importance. HER2 status is usually determined by immunohistochemistry (IHC) and occasionally by *in situ* hybridization (ISH). However, little is known regarding the performance difference among the commercially available anti-HER2 immunohistochemical antibodies in gastric adenocarcinomas. This study compared three anti-HER2 antibodies (HercepTest, 4B5 and SP3) using two different arms: samples prepared in tissue microarray (TMA) device and whole-tissue sections counterpart prepared from paraffin blocks of a series of 199 gastric adenocarcinomas.

Research frontiers

HER2 is a member of the family of tyrosine kinase receptors. Overexpression of the HER2 receptor has been identified in various cancers and is most widely studied in breast cancer. Like in breast cancer, HER2-positive gastric tumors are correlated with worse prognosis. An important difference in HER2 immunoreactivity between breast cancer and gastric adenocarcinomas is the striking heterogeneity of HER2-positivity in the latter, which can affect the determination of HER2 status in small samples such as biopsies and TMA.

Innovations and breakthroughs

Many studies have compared anti-HER2 antibodies in breast cancer, however only few reports have done it in gastric cancer. Gastric adenocarcinomas exhibit intratumoral heterogeneity of HER2 expression and have a unique IHC scoring system. This is the first study to compare HercepTest, 4B5 and SP3 in a series of gastric tumors. Most contemporary studies use the cost-effective TMA technique for testing HER2 expression; however, given the high incidence of heterogeneous HER2-immunoreactivity and the risk of underestimating the incidence of HER2-amplification rate; we also aimed to compare the results obtained from TMA to those obtained from whole-tissue sections.

Applications

The study results indicate the best method to assess HER2 expression in gastric cancer, facilitating the choice of the antibody with the greatest ability to identify the most patients who could benefit from the use of trastuzumab. The results also demonstrated that HER2 expression in small samples of gastric cancer should be cautiously evaluated because the intratumoral heterogeneity may influence the results.

Terminology

According to the four-tiered IHC scoring system for gastric cancer, samples scored as 0 and 1+ are negative, 2+ as equivocal and 3+ as positive. Intratumoral heterogeneity is defined as areas with different HER2 scores within the

same tumor.

Peer review

This is a highly stringent study in which the authors examined HER2 immunostaining in a series of gastric cancers, using three different commercially available anti-HER2 antibodies in samples of TMA and whole-tissue sections. The sensitivity, predictive value for HER2 amplification and accuracy of each antibody were determined. The results are provoking and indicate the most accurate IHC method for HER2 evaluation, emphasizing the limitations of the TMA technique. The results of the study should encourage a more judicious assessment of the HER2 status in gastric cancers, especially in small tissue samples.

REFERENCES

- Dicken BJ, Bigam DL, Cass C, Mackey JR, Joy AA, Hamilton SM. Gastric adenocarcinoma: review and considerations for future directions. *Ann Surg* 2005; **241**: 27-39 [PMID: 15621988]
- INCA, 2012. Available from: URL: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/estomago>
- Jackson C, Cunningham D, Oliveira J. Gastric cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol* 2009; **20** Suppl 4: 34-36 [PMID: 19454457 DOI: 10.1093/annonc/mdp122]
- Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006; **24**: 2137-2150 [PMID: 16682732 DOI: 10.1200/JCO.2005.05.2308]
- Akiyama T, Sudo C, Ogawara H, Toyoshima K, Yamamoto T. The product of the human c-erbB-2 gene: a 185-kilodalton glycoprotein with tyrosine kinase activity. *Science* 1986; **232**: 1644-1646 [PMID: 3012781]
- Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987; **235**: 177-182 [PMID: 3798106]
- Allgayer H, Babic R, Gruetzner KU, Tarabichi A, Schildberg FW, Heiss MM. c-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. *J Clin Oncol* 2000; **18**: 2201-2209 [PMID: 10829039]
- Grabsch H, Sivakumar S, Gray S, Gabbert HE, Müller W. HER2 expression in gastric cancer: Rare, heterogeneous and of no prognostic value - conclusions from 924 cases of two independent series. *Cell Oncol* 2010; **32**: 57-65 [PMID: 20208134 DOI: 10.3233/CLO-2009-0497]
- Hofmann M, Stoss O, Shi D, Büttner R, van de Vijver M, Kim W, Ochiai A, Rüschoff J, Henkel T. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. *Histopathology* 2008; **52**: 797-805 [PMID: 18422971 DOI: 10.1111/j.1365-2559.2008.03028.x]
- Jørgensen JT. Targeted HER2 treatment in advanced gastric cancer. *Oncology* 2010; **78**: 26-33 [PMID: 20185938 DOI: 10.1159/000288295]
- Kim KC, Koh YW, Chang HM, Kim TH, Yook JH, Kim BS, Jang SJ, Park YS. Evaluation of HER2 protein expression in gastric carcinomas: comparative analysis of 1,414 cases of whole-tissue sections and 595 cases of tissue microarrays. *Ann Surg Oncol* 2011; **18**: 2833-2840 [PMID: 21468783 DOI: 10.1245/s10434-011-1695-2]
- Park DI, Yun JW, Park JH, Oh SJ, Kim HJ, Cho YK, Sohn CI, Jeon WK, Kim BI, Yoo CH, Son BH, Cho EY, Chae SW, Kim EJ, Sohn JH, Ryu SH, Sepulveda AR. HER-2/neu amplification is an independent prognostic factor in gastric cancer. *Dig Dis Sci* 2006; **51**: 1371-1379 [PMID: 16868827 DOI: 10.1007/s10620-005-9057-1]
- Tanner M, Hollmén M, Junttila TT, Kapanen AI, Tommola S, Soini Y, Helin H, Salo J, Joensuu H, Sihvo E, Elenius K, Isola

- J. Amplification of HER-2 in gastric carcinoma: association with Topoisomerase IIalpha gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. *Ann Oncol* 2005; **16**: 273-278 [PMID: 15668283 DOI: 10.1093/annonc/mdi064]
- 14 **Wang S**, Zheng G, Chen L, Xiong B. Effect of HER-2/neu over-expression on prognosis in gastric cancer: a meta-analysis. *Asian Pac J Cancer Prev* 2011; **12**: 1417-1423 [PMID: 22229170]
- 15 **Bang YJ**, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, Hill J, Lehle M, Rüschoff J, Kang YK. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010; **376**: 687-697 [PMID: 20728210 DOI: 10.1016/S0140-6736(10)61121-X]
- 16 **Hudis CA**. Trastuzumab--mechanism of action and use in clinical practice. *N Engl J Med* 2007; **357**: 39-51 [PMID: 17611206 DOI: 10.1056/NEJMra043186]
- 17 National Cancer Institute. Targeted Cancer Therapies (NCI Web site). Available from: URL: <http://www.cancer.gov/cancertopics/factsheet/Therapy/targeted>
- 18 **Boers JE**, Meeuwissen H, Methorst N. HER2 status in gastro-oesophageal adenocarcinomas assessed by two rabbit monoclonal antibodies (SP3 and 4B5) and two in situ hybridization methods (FISH and SISH). *Histopathology* 2011; **58**: 383-394 [PMID: 21323962 DOI: 10.1111/j.1365-2559.2011.03760.x]
- 19 **Hoang MP**, Sahin AA, Ordóñez NG, Sneige N. HER-2/neu gene amplification compared with HER-2/neu protein over-expression and interobserver reproducibility in invasive breast carcinoma. *Am J Clin Pathol* 2000; **113**: 852-859 [PMID: 10874886 DOI: 10.1309/VACP-VLQA-G9DX-VUDF]
- 20 **Manion E**, Hornick JL, Lester SC, Brock JE. A comparison of equivocal immunohistochemical results with anti-HER2/neu antibodies A0485 and SP3 with corresponding FISH results in routine clinical practice. *Am J Clin Pathol* 2011; **135**: 845-851 [PMID: 21571957 DOI: 10.1309/AJCP5L003NGDJG]
- 21 **Yano T**, Doi T, Ohtsu A, Boku N, Hashizume K, Nakanishi M, Ochiai A. Comparison of HER2 gene amplification assessed by fluorescence in situ hybridization and HER2 protein expression assessed by immunohistochemistry in gastric cancer. *Oncol Rep* 2006; **15**: 65-71 [PMID: 16328035]
- 22 **Dietel M**, Ellis IO, Höfler H, Kreipe H, Moch H, Dankof A, Kölbl K, Kristiansen G. Comparison of automated silver enhanced in situ hybridisation (SISH) and fluorescence ISH (FISH) for the validation of HER2 gene status in breast carcinoma according to the guidelines of the American Society of Clinical Oncology and the College of American Pathologists. *Virchows Arch* 2007; **451**: 19-25 [PMID: 17562074 DOI: 10.1007/s00428-007-0424-5]
- 23 **Koh YW**, Lee HJ, Lee JW, Kang J, Gong G. Dual-color silver-enhanced in situ hybridization for assessing HER2 gene amplification in breast cancer. *Mod Pathol* 2011; **24**: 794-800 [PMID: 21317877 DOI: 10.1038/modpathol.2011.9]
- 24 **Yang J**, Luo H, Li Y, Li J, Cai Z, Su X, Dai D, Du W, Chen T, Chen M. Intratumoral heterogeneity determines discordant results of diagnostic tests for human epidermal growth factor receptor (HER) 2 in gastric cancer specimens. *Cell Biochem Biophys* 2012; **62**: 221-228 [PMID: 21927816 DOI: 10.1007/s12013-011-9286-1]
- 25 **Ricardo SA**, Milanezi F, Carvalho ST, Leitão DR, Schmitt FC. HER2 evaluation using the novel rabbit monoclonal antibody SP3 and CISH in tissue microarrays of invasive breast carcinomas. *J Clin Pathol* 2007; **60**: 1001-1005 [PMID: 17158643 DOI: 10.1136/jcp.2006.040287]
- 26 **Torrisi R**, Rotmensz N, Bagnardi V, Viale G, Curto BD, Dell'orto P, Veronesi P, Luini A, D'Alessandro C, Cardillo A, Goldhirsch A, Colleoni M. HER2 status in early breast cancer: relevance of cell staining patterns, gene amplification and polysomy 17. *Eur J Cancer* 2007; **43**: 2339-2344 [PMID: 17855075 DOI: 10.1016/j.ejca.2007.07.033]
- 27 **Cohen J**. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; **70**: 213-220 [PMID: 19673146]
- 28 **Cho EY**, Srivastava A, Park K, Kim J, Lee MH, Do I, Lee J, Kim KM, Sohn TS, Kang WK, Kim S. Comparison of four immunohistochemical tests and FISH for measuring HER2 expression in gastric carcinomas. *Pathology* 2012; **44**: 216-220 [PMID: 22437741 DOI: 10.1097/PAT.0b013e3283513e8b]
- 29 **Dekker TJ**, Borg ST, Hooijer GK, Meijer SL, Wesseling J, Boers JE, Schuurung E, Bart J, van Gorp J, Mesker WE, Kroep JR, Smit VT, van de Vijver MJ. Determining sensitivity and specificity of HER2 testing in breast cancer using a tissue micro-array approach. *Breast Cancer Res* 2012; **14**: R93 [PMID: 22694844 DOI: 10.1186/bcr3208]
- 30 **Thomson TA**, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, Dupuis B, Parker RL. HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. *Mod Pathol* 2001; **14**: 1079-1086 [PMID: 11706067 DOI: 10.1038/modpathol.3880440]
- 31 **Drev P**, Grazio SF, Bracko M. Tissue microarrays for routine diagnostic assessment of HER2 status in breast carcinoma. *Appl Immunohistochem Mol Morphol* 2008; **16**: 179-184 [PMID: 18227723 DOI: 10.1097/PAI.0b013e31806875e1]
- 32 **Kim MA**, Lee HJ, Yang HK, Bang YJ, Kim WH. Heterogeneous amplification of ERBB2 in primary lesions is responsible for the discordant ERBB2 status of primary and metastatic lesions in gastric carcinoma. *Histopathology* 2011; **59**: 822-831 [PMID: 22092393 DOI: 10.1111/j.1365-2559.2011.04012.x]
- 33 **Perrone G**, Amato M, Callea M, Rabitti C, Righi D, Crucitti P, Coppola R, Onetti Muda A. HER2 amplification status in gastric and gastro-oesophageal junction cancer in routine clinical practice: which sample should be used? *Histopathology* 2012; **61**: 134-135 [PMID: 22551459 DOI: 10.1111/j.1365-2559.2012.04251.x]

P- Reviewer Piscaglia AC S- Editor Wen LL L- Editor A
E- Editor Zhang DN

