

HER2 diagnostics in gastric cancer—guideline validation and development of standardized immunohistochemical testing

Josef Rüschoff · Manfred Dietel · Gustavo Baretton · Susanne Arbogast · Axel Walch · Geneviève Monges · Marie-Pierre Chenard · Frédérique Penault-Llorca · Iris Nagelmeier · Werner Schlake · H. Höfler · H. H. Kreipe

Received: 10 November 2009 / Revised: 2 July 2010 / Accepted: 13 July 2010 / Published online: 28 July 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Trastuzumab-based therapy has been shown to confer overall survival benefit in HER2-positive patients with advanced gastric cancer in a large multicentric trial (ToGA study). Subgroup analysis identified adenocarcinomas of the stomach and gastroesophageal (GE) junction with overexpression of HER2 according to immunohistochemistry (IHC) as potential responders. Due to recent approval of trastuzumab for HER2 positive metastatic gastric and GE-junction cancer in Europe (EMEA) HER2 diagnostics is now mandatory with IHC being the primary test followed by fluorescence in situ hybridization (FISH) in IHC2+ cases. However, in order to not miss patients potentially responding to targeted therapy determination of

a HER2-positive status for gastric cancer required modification of scoring as had been proposed in a pre-ToGA study. To validate this new HER2 status testing procedure in terms of inter-laboratory and inter-observer consensus for IHC scoring a series of 547 gastric cancer tissue samples on a tissue microarray (TMA) was used. In the first step, 30 representative cores were used to identify specific IHC HER2 scoring issues among eight French and German laboratories, while in the second step the full set of 547 cores was used to determine IHC HER2 intensity and area score concordance between six German pathologists. Specific issues relating to discordance were identified and recommendations formulated which proved to be effective

J. Rüschoff (✉) · I. Nagelmeier
Targos Molecular Pathology GmbH und Pathologie Nordhessen,
Germaniastr 7,
34119 Kassel, Germany
e-mail: rueschoff@patho-nordhessen.de

M. Dietel
Department of Pathology, Humboldt University, Campus Charité,
Berlin, Germany

G. Baretton
Institute of Pathology,
University Hospital Carl Gustav Carus Dresden,
Dresden, Germany

S. Arbogast
Department of Pathology & Tissue Biomarkers, MML,
Pharma Development, Roche Diagnostics GmbH,
Penzberg, Germany

A. Walch
Institute of Pathology, Helmholtz Zentrum München—German
Research Center for Environmental Health (GmbH),
Neuherberg, Germany

G. Monges
Département de Biopathologie, Institut Paoli Calmettes,
Marseille, France

M.-P. Chenard
Département de Pathologie, Hôpital de Hautepierre,
Strasbourg, France

F. Penault-Llorca
Service d'Anatomo-Pathologie, Centre Jean Perrin,
Clermont-Ferrand, France

W. Schlake
Pathologisches und Gewerbepathologisches Institut,
Gelsenkirchen, Germany

H. Höfler
Pathological Institute, University of Munich,
Munich, Germany

H. H. Kreipe
Institute of Pathology, Medizinische Hochschule Hannover,
Hannover, Germany

to reliably determine HER2 status in a prospective test series of 447 diagnostic gastric cancer specimens.

Keywords HER2 · Immunohistochemistry (IHC)
Gastric carcinoma · Diagnostic test

Introduction

The human epidermal growth receptor 2 gene (*HER2*, also known as *ERBB2* and *HER2/neu*) is now well recognized as a key in the development of certain solid human tumors, most notably in breast cancer. In breast cancer, *HER2* gene amplification almost invariably induces and occurs before HER2 protein overexpression on the tumor cell surface [1, 2]. Monitoring of the tumor HER2 status (gene amplification and/or protein overexpression) in breast cancer has become routine as the positive HER2 status detected in around 25% of these cancers is associated with poorer prognosis, more aggressive disease, and an increased risk of disease recurrence [3–5]. Furthermore, in breast cancer the determination of HER2 status is necessary for optimal application of HER2-directed therapies such as trastuzumab (Herceptin®, Roche), which increases overall survival in both the metastatic [6, 7] and adjuvant settings [8–10], and predicts response in the neo-adjuvant setting [9, 10].

Data reported in the literature for HER2 positivity rates in gastric cancer vary from about 7–43% (for review see, [11]) with most studies demonstrating values of about 15–25% [11–14]. Furthermore, a HER2-positive status in gastric cancer also appears to be associated with poorer prognosis, more aggressive disease, and shorter survival [12, 14–19]. Preclinical studies have indicated that trastuzumab exerts antitumor activity in HER2-overexpressing human gastric cell lines and xenograft models [19–21]. As a consequence, the addition of trastuzumab to fluoropyrimidine/platinum-based therapy has been investigated in a large-scale ($n=584$) randomized study in patients with HER2-positive advanced gastric cancer, the ToGA trial, the primary results of which have been presented and showed that trastuzumab significantly improved the primary endpoint, median overall survival, by nearly 3 months (11.1 to 13.8 months, $p=0.0046$) with no impact on overall treatment safety. Moreover, an increased benefit from trastuzumab treatment was seen for patients who had higher levels of HER2 protein expression, including subgroups for IHC2+/FISH+ and IHC3+ (median survival increased from 11.8 months for the chemotherapy treatment arm to 16.0 months for the chemotherapy with trastuzumab arm) [22, 23]. Thus, very recently the European board, EMEA, approved trastuzumab for the treatment of metastasized adenocarcinomas of the stomach and the esophageal junction [24]. Thereby immunohistochemical testing is the

primary method of choice to determine HER2 status in gastric cancer. FISH is restricted to those cases that have equivocal (IHC2+) HER2 expression.

The method of HER2 scoring within the phase III ToGA trial was essentially based on a separate (so-called pre-ToGA) validation study where immunohistochemistry (IHC) protein expression and fluorescence in situ hybridization (FISH) gene amplification was correlated in a series of 168 gastric cancer resection specimens [11]. An international consensus was reached to modify the breast scoring system for IHC by accepting strong incomplete (basolateral) membranous staining as positive (3+) and by abolishing the 10% area cut-off for this group in biopsies. A patient was considered to have HER2-positive gastric cancer with a score of IHC3+ (HerceptTest) and/or FISH-positive result. Screening of nearly 4,000 patients in 24 countries for entry to the ToGA trial revealed a 22.1% HER2-positivity rate [25]. Furthermore, HER2-positivity rates were higher in gastroesophageal junction than gastric cancer (33% vs. 21%, $p<0.001$) and in intestinal than diffuse or mixed cancer (32.2% vs. 6.1% vs. 20.4%, $p<0.001$). Concordance between IHC and FISH was 87.5%: while in breast cancer, most IHC0/1+ results are FISH negative, the frequency of IHC0/1+ samples testing as FISH-positive was similar as IHC2+/FISH-positive samples (23% vs. 26%) [25].

The aim of our current study was to validate this HER2 testing procedure by determining whether pathologists from different sites are able to reproduce the method of gastric cancer HER2 status evaluation as it was used by the study pathologist (JR) within the ToGA study.

In a first step, inter-laboratory variation was assessed using 30 gastric cancer core biopsy tissue microassay samples (TMAs) using different HER2 IHC methods. In a second step, inter-observer variation of HER2 IHC scoring was tested in a series of already stained gastric cancer TMAs ($n=547$). A consensus practical guideline for accurate HER2 analysis in gastric cancer was then developed on the basis of these data. Finally, for validation of these guidelines a series of 447 prospective diagnostic gastric cancer specimens were tested at five participating sites throughout Germany.

Materials and methods

All study TMAs were based on a series of 547 gastric cancer core biopsy specimens assembled and provided by Prof. Kreipe (Hannover). All individual patients gave written informed consent for biological studies at their initial presentation. All samples were obtained from surgery performed for diagnostic and/or therapeutic purposes and were used according to German ethical regulations. The study followed the guidelines of the Declaration of Helsinki and

patient identity of the pathological specimens remained anonymous in the context of this study.

In a first step, a 30-core TMA set was pre-selected that provided a representative of all tumor types (intestinal, mixed, and diffuse) according to Lauren classification as well as different HER2 expression and amplification levels. Core size was 0.4 cm and each represented a different tumor sample. HER2 assessment was performed using different commercial assays according to the manufacturers' instructions at the different participating sites ($n=8$). IHC immunostaining was conducted using HercepTest™ (Dako Denmark A/S, Glostrup, Denmark) and/or the PATHWAY® HER2/neu (4B5) antibody (Ventana Medical Systems SA, Illkirch, France). *HER2* amplification was determined by FISH assays, using either *HER2* FISH pharmDX™ (Dako Denmark A/S) or PathVysion® (Abbott Laboratories, Des Plaines, IL, USA). Automated bright-field dual-color silver in situ hybridization (SISH) assay (BDISH; Inform™, Ventana Medical Systems SA) was used to determine gene amplification at three of the participating sites [26]. Evaluation was performed according to the modified gastric cancer testing protocol [11] taking incomplete basolateral or only lateral staining into account. As TMA cores were tested analogous to biopsies the 10% cut-off was recorded but not regarded for the final scoring (i.e., 1+, 2+, and 3+). FISH and BDISH were performed according to the manufacturers' recommendations with ratios above 2.0 being considered amplified.

In a second step, the complete TMA sample series of 547 tumor cores was used to determine inter-observer variation of HER2 expression (staining intensity and area stained) scoring independent of inter-laboratory staining variation. Thus, TMAs used for evaluation were already IHC stained using the 4B5 antibody (Ventana Medical Systems SA) at the Hannover laboratory that supplied samples.

Data for the first 30-core TMA set were presented and discussed at a 2-day consensus meeting conducted at the Institute of Pathology, Charite, Berlin (27/28 March 2009). After a consensus was reached about specific issues concerning HER2 scoring by IHC in gastric cancer, the second full set of 547 cores were evaluated independently by six German pathologists (GB, MD, HH, JR, SA, AW). The complete 547 TMA set was scanned (Provito GmbH, Berlin, Germany) and provided as virtual slides to the panelists. By use of this data set, all cases that resulted in discordant IHC scores between observers were then individually discussed at a separate meeting in Düsseldorf (10 June 2009) to determine the most reproducible practical guideline for HER2 testing in gastric cancer. Statistical analyses were performed using the statistical program R version 1.9.1 and Microsoft® Excel®. Kappa statistics was calculated according to the method of Conger (1980) by package “irr” of program R [33]. In order validate

these guidelines in routine practice a series of $n=447$ prospective diagnostic gastric cancer samples have been tested at five different participating sites throughout Germany which comprised either a biopsy tissue block or one representative tissue block of resection specimen. Thereby, four sites followed the algorithm as proposed by EMEA with IHC (4B5, Ventana) being used first and one center applied both IHC and ISH (BDISH, Ventana) to all $n=152$ specimens at their site.

Results

Inter-laboratory reproducibility of *HER2* scoring in gastric cancer

In total, 29 of 30 cores of the first TMA set showed evaluable tumor tissue when evaluated by HercepTest and 4B5 antibody at 7 and 8 sites, respectively. The core with non-evaluable tumor tissue was excluded. A HER2 score deviation ≤ 1 was found in 14/29 cores (48.3%) when HercepTest was used as compared to 22/29 cores (75.9%, $p=0.002$) when 4B5 was used, despite one more site being included for the latter. Consensus HER2 scores was reached for all but 1 site for 11 (37.9%) tumors with HercepTest and by 13 (44.8%) tumors with 4B5. Consensus HER2 scores between sites with each IHC test increased as a function of fewer sites being in agreement.

In total, 27 tumor cores were evaluable for *HER2* gene amplification by hybridization assays. Comparison with *HER2* amplification according to FISH/BDISH results with HER2 IHC scores unanimously agreed by all sites as negative (0/1+) or positive (2+/3+) revealed a tendency towards higher sensitivity for 4B5 detecting positive *HER2* amplification (Table 1). In particular, five out of eight ISH-positive tumor cores were scored as 2+/3+ by 4B5 whereas this was the case in only two cores with HercepTest; IHC scoring was equivocal with some positive and some negative classifications in three cases by 4B5 and six cases by HercepTest, respectively. There was no difference between both test platforms with respect to the ISH-negative cases with the one equivocal case testing negative at some sites and as positive at others.

Pitfalls and rules of *HER2* IHC scoring

Reviewing of the 30 core TMA slides disclosed several pitfalls that were mainly related to interpretation of staining rather than staining variation at different lab sites. After panel discussion the following pitfalls turned out to be the major reasons for inter-observer variation. Figure 1 shows examples of potential pitfalls and of different scores that can be obtained by HER2 receptor staining. HER2

Table 1 Inter-laboratory comparison: comparison of *HER2* amplification status according to FISH/BDISH results with *HER2* IHC scores

FISH/BDISH result (<i>n</i> =27)	IHC score					
	4B5			HercepTest		
	Negative (0/1+)	Equivocal (0/1+ vs. 2+/3+)	Positive (2+/3+)	Negative (0/1+)	Equivocal (0/1+ vs. 2+/3+)	Positive (2+/3+)
Negative (<i>n</i> =18)	12	5	1	11	7	0
Equivocal (<i>n</i> =1)	0	1	0	0	1	0
Positive (<i>n</i> =8)	0	3	5	0	6	2

expression may occur in areas of gastric mucosal metaplasia and towards reactive epithelial cells bordering ulcers (Fig. 1a). Discordant results can also occur where <10% of tumor cells are stained. This holds particularly true if only a few (<5) cells were evaluated. Such small cell groups tend to show unspecific rather pericellular and granular instead of distinct intercellular staining particularly at the edges of the TMA cores (Fig. 1b). Another source of false positive scoring is diffuse cytoplasmic reaction with or without nuclear staining (Fig. 1c). For evaluation of membranous staining specific for scoring it turned out that consideration of microscope magnification is of importance to reach high inter-observer agreement. Accordingly, strong *HER2*-positive (3+) staining may be visible to the naked eye and displays unequivocal membranous staining already at low magnification ($\times 2.5/\times 5$) as shown in Fig. 2d,e. An example of heterogeneous staining is given in Fig. 1f. Areas of 3+ *HER2*-positive cells (visible at low magnification $\times 2.5$) in less than 10% of tumor cells are admixed with those where unequivocal membranous staining can only be disclosed at medium magnification (first at $\times 10$ corresponding to 2+ intensity) and with some areas showing barely visible expression (membranous staining confirmed only at $\times 40$, corresponding to 1+). Fig. 1g shows focal staining (<10% of the tumor) where unequivocal membranous staining could first be recognized at medium magnification ($\times 10$, Fig. 1h). In contrast demonstration of IHC 1+ score needs high magnification ($\times 40$) to confirm an unequivocally intercellular membranous *HER2* expression (Fig 1i).

Figure 2 summarizes the recommended stepwise process of standardized IHC scoring in gastric cancer. At first tumor and tissue quality issues have critically to be checked. Most of all, unspecific staining within non-neoplastic lesions such as intestinal metaplasia have to be excluded from scoring and also edge and crushing artifacts affecting tumor cells. In a second step, only distinct membranous staining either complete (chicken-wire type), basolateral or only lateral between cell–cell contacts in a cluster of at least five cells are considered if biopsies are scored. This excludes any equivocal staining, e.g., rather thickened or granular at basal cell basement membrane or at a single isolated cell

surrounded by a shrinkage rim. Any staining at the luminal site of a tumor gland is suspicious of artifactual staining not specific for scoring particularly if it is not associated with distinct intercellular staining and should therefore be excluded. In the third step the final scoring should be done taking the microscope magnification into account at which unequivocal membrane staining can readily be confirmed.

Inter-observer reproducibility of *HER2* scoring in gastric cancer

After discussion and definition of the above-mentioned pitfalls and rules for reliable IHC scoring the full TMA set (547 cores) stained by 4B5 was graded with respect to both intensity and area scores by six pathologists. A *HER2* score deviation ≤ 1 between pathologists was now found for 95.6% and 91.8% of TMAs for intensity and area scores, respectively. Consensus agreement for *HER2* staining intensity scores between all except 1 or 2 of the 6 pathologists was found for 512/547 (93.6%) and 470/547 (85.9%) tumors, respectively, while corresponding consensus agreement for area stained scores rates were 505/547 (92.3%) and 465/547 (85.0%) tumors, respectively.

Calculation of interrater kappa values showed moderate agreement ($k=0.61$) if all intensity scores (IHC 0–3+) were considered. However, if negative (IHC 0/1+) was calculated against positive (IHC 2+/3+) kappa value rose to 0.805 indicating an “almost perfect agreement”. In fact, this implies that 92.5% of the six pathologists were in complete agreement in all 547 cores. In five cases, there was marked discordance with half scoring a case IHC 0/1+ and half IHC 2+/3+. Reevaluation of these cases (Düsseldorf meeting) disclosed that in all of these cases the number of evaluable tumor cells was below five cells.

Validation of guidelines by application to diagnostic specimens

In order to demonstrate the value of the defined IHC scoring rules with respect to routine practice and gene amplification status a total of 447 diagnostic gastric cancer

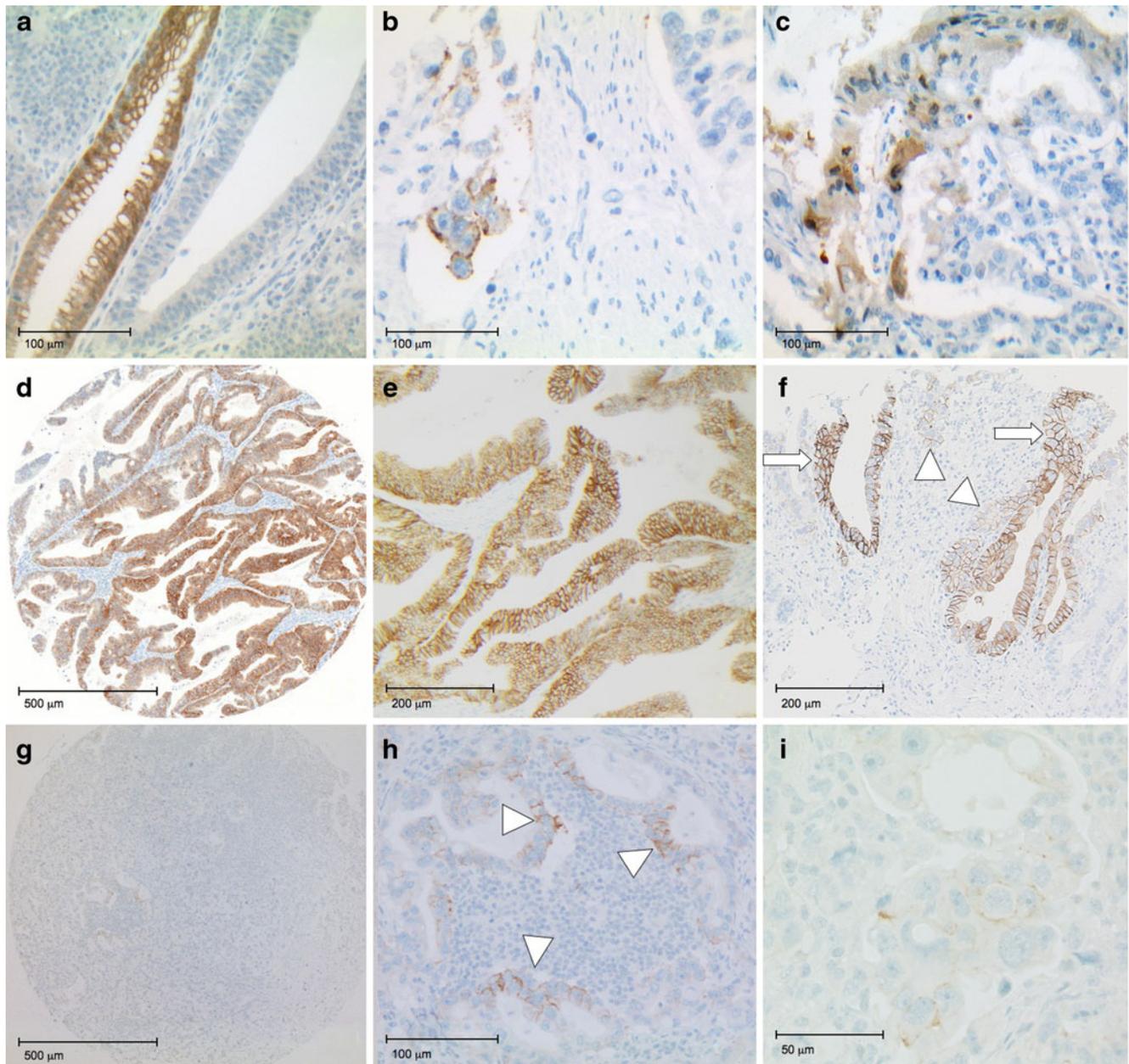


Fig. 1 Photomicrographs of TMA examples. **a–c** Artifacts leading to potential mis-scoring on IHC: **a** intestinal metaplasia, **b** edge artifact at TMA border with granular (not linear) pseudo-membranous staining, and **c** cytoplasmic as well as nuclear staining. **d–h** Intensity scoring: **d** Score 3+ visible by naked eye with membranous staining clearly visible at low magnification (obj. $\times 5$) being either complete, basolateral or lateral (**e**, $\times 10$). **f** Photomicrograph of TMA sample showing distinction between 2+ and 3+ IHC using 4B5 antibody. *Arrows* indicate areas with clearly visible membrane staining at low

magnification (i.e., 3+), focally in <10% of tumor; *arrowheads* indicate areas where membrane staining is only visible at $\times 10$ magnification (i.e., 2+). **g** TMA core suspicious of some focal staining at $\times 5$ which turned out to be a focally specific membranous staining in groups of at least five cells at medium magnification (**h**, $\times 20$; see *arrowheads*). **i** Very weak staining where membranous staining is barely visible and could only be demonstrated using high magnification (**i**, $\times 40$)

specimens were prospectively tested at five participating sites (Berlin, Dresden, Hannover, Kassel, Munich). The mean positivity rate was 22.8% (102/447). In one site, both IHC and ISH test were applied in parallel to a total of 152 gastric carcinomas. Accordingly, the consented testing guidelines resulted in a high concordance between IHC

and ISH. All IHC3+ tumors ($n=24$) showed *HER2* gene amplification (100% concordance), which was the case in 32% of the IHC2+ ($n=47$) and in 5% of the IHC 1+ ($n=41$) cases. Interestingly, most of IHC2+ and IHC1+ (76.5%) showed low level amplification (ratio 2–3) which was only the case in 16% of IHC3+ tumors.

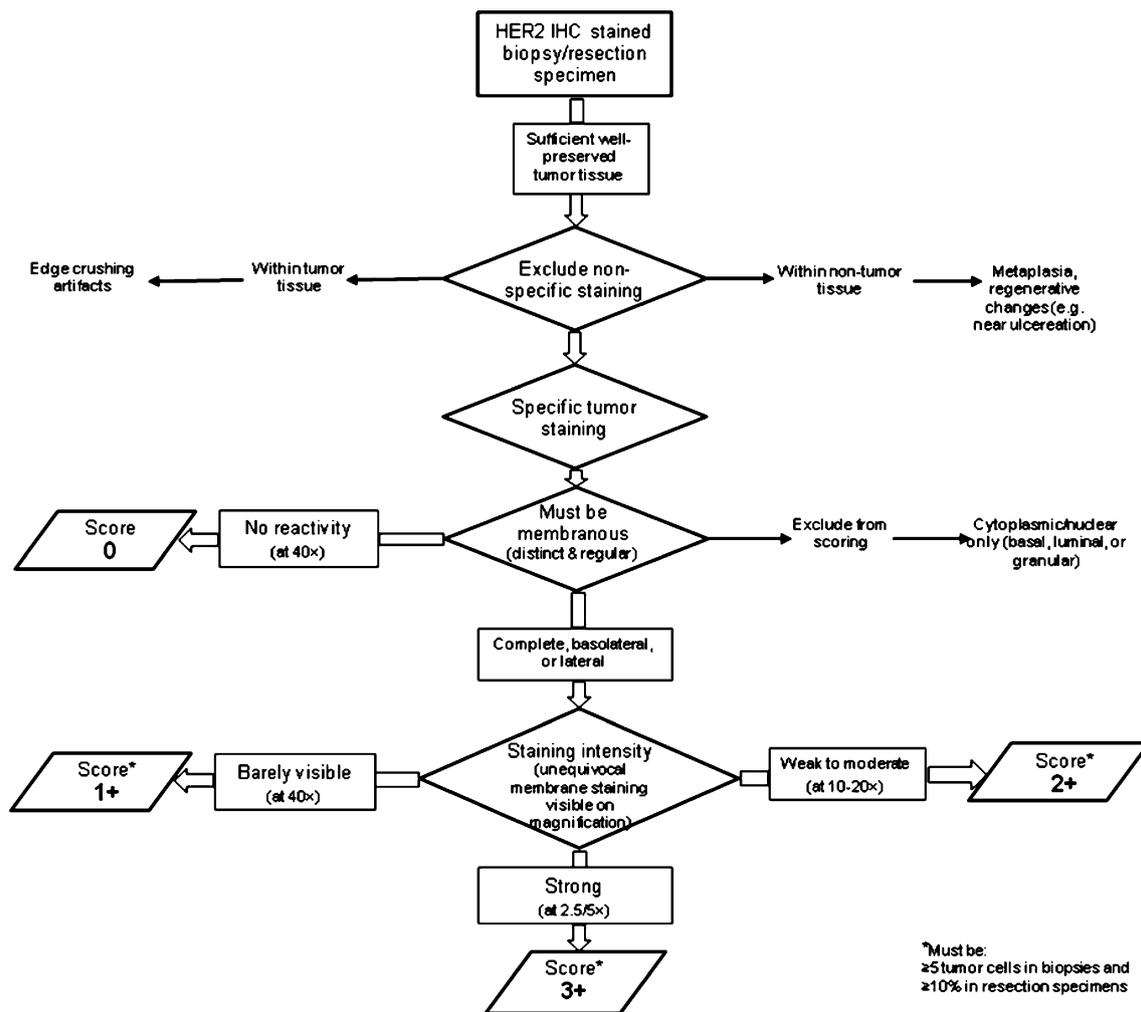


Fig. 2 Stepwise approach to IHC scoring in gastric cancer: tissue and quality issues (mod. acc. to [31])

Discussion

The modified HER2 immunoscore method that has been shown to be predictive of response to trastuzumab-based therapy in patients with advanced gastric cancer in the ToGA study [22] was shown to be reproducible between different pathologists in our study as long as special precautions were adopted. Furthermore, our study indicates that HER2 status determination in gastric cancer needs special and specific training and guidelines for pathologists in a similar manner to that already adopted for pathologists for breast cancer testing. However, the specific guidelines for breast cancer are not applicable to gastric cancer in several aspects, as there are important and significant differences in HER2 status determination between breast and gastric cancer (Table 2).

Thus determination of HER2 status by just transferring the breast cancer IHC scoring roles to gastric cancer may lead to a significant loss of patients. In a recent paper by Barros-Silva et al. [27], resection specimens of 463 gastric adenocarcinomas were tested using the breast cancer

scoring rules. Accordingly, these authors classified 3.9% as IHC2+ and 5.4% as IHC3+. The corresponding values in the TOGA trial with 12% IHC2+ and 11% IHC3+ were about twice as high [23]. The same holds true if one compares TMA data which were classified as IHC2+ (1.6%) or IHC3+ (3.2%) if breast cancer scoring was applied [28]. As the same group also tested gastric cancer TMAs using our proposed gastric cancer specific scoring [11] the corresponding rates were 4% IHC2+ and 13% for IHC3+, demonstrating an about fourfold increase of HER2 positivity rate [29]. Therefore, it is supposed that application of breast cancer scoring to gastric cancer may produce an up to 50% false-negative rate if IHC is used as the primary test platform as favored by EMEA.

On the other hand, if FISH is used as first screening step only few IHC3+ positive cases may be missed but one may be faced with a quite high percentage of non-responders according to ToGA data [22, 23].

Therefore, based on the presented consensus study a number of specific recommendations can be made for reliable

Table 2 HER2 diagnostics in gastric cancer—differences to breast cancer (acc. to [32])

		Gastric cancer	Breast cancer
IHC scoring	Extent (Area Cut-off)	Biopsy specimens ≥ 5 Cells Resection specimens: $\geq 10\%$	$\geq 10\%$ ($\geq 30\%$) ^a
	Circularity	Mostly missing (often only lateral in IHC2+/3+)	A must in IHC2+/3+
(F)ISH analysis	Cell number	20 cohesive tumor cells showing highest gene count (add 20 new if ratio 1.8–2.2)	20 cohesive tumor cells showing highest gene count (add 20 new if ratio 1.8–2.2)
HER2 positivity	Amplification	Ratio ≥ 2.0	Ratio ≥ 2.0 (≥ 2.2) ^a
	Tumor type	About 30% of intestinal-type GC about 15% of mixed-type about 5% of diffuse type (signet ring type typically negative)	15–25% of ductal type (G2/G3); almost never in subtypes such as lobular, medullar and ductal G1
	Tumor location	About 30% at cardiac/GE-junction about 15% of gastric cancer	No correlation
Patient selection	FISH vs. IHC	IHC more predictive than FISH: IHC primary test ^b FISH only if IHC2+	FISH/IHC equally predictive:

^a Acc. to ASCO/CAP [31]

^b Acc. to approval by EMEA

HER2 status determination in gastric cancer. Since HER2 expression is mainly restricted to intestinal-type, gland-forming gastric cancer cells incomplete, often basolateral or only lateral, membranous IHC staining is the rule rather than an exception for HER2-positive gastric cancer [11]. Thus, unlike for breast cancer, circularity of IHC staining is no longer a criterion for HER2 IHC scoring in gastric cancer.

A key issue for the classification of positive HER2 expression is a membranous staining that can be unequivocally assessed as linear staining at cell–cell contact sites. Strong tumor HER2 IHC staining is usually already directly visible, particularly if most of the tumor is stained. In these cases, only low magnification ($\times 2.5$ – 5) is needed to confirm strong staining intensity. In any case where high magnification ($\times 40$) is required for unequivocal demonstration of membranous staining, the tumor is scored IHC 1+ (Fig. 1i). It should, however, be mentioned that quality of lenses and brightness of the microscopic light might be of influence as well. Nevertheless the inter-observer variation results within this ring-study prior and after application of the magnification rule clearly are in favor of such an approach over non-standardized wording, e.g., of “barely visible” for IHC1+.

Recommendation 1

For reproducible intensity scoring, the degree of microscopic magnification (x -fold) at which membranous (linear intercellular) staining is clearly visible should be considered.

According to the pre-ToGA HER2 validation study [11], IHC HER2 expression is often focal in gastric cancer which could largely be confirmed by the ToGA study [25]. An

assessment area cut-off of at least 10% stained tumor cells, as originally proposed for HER2 scoring in breast cancer, was omitted in gastric cancer biopsies. A rather focal HER2 staining was also frequently observed in our current series of TMAs. Inter-observer agreement was especially low if less than five cells were stained. The minimum number of cells that could reliably be assessed was five. Thus, in biopsies a focus (“clone”) that is allowed to be scored should have at least five stained evaluable cells.

Recommendation 2

Since focal HER2 expression is an issue in gastric cancer, biopsies should only be evaluated if at least five cohesive, unequivocal tumor cells are stained. In resection specimens, the 10% cut-off should be kept. Thus scoring procedure is different for biopsies and resection specimen [23].

Thereby, only unequivocal intercellular staining is accepted; even ring-shaped staining of a single tumor cell is not accepted either due to the cell number criterion (see above) or difficulties in exclusion of edge artifacts. The pre-ToGA HER2 scoring validation study used HercepTest as the only IHC assay [11]. The different laboratory sites generally used two different IHC assays (HercepTest and 4B5) concurrently. Both assays resulted in a somewhat degree high inter-laboratory discordance, which appeared to be higher when HercepTest was used. These observations could essentially be confirmed by the validation set where 4B5 was used in 447 prospective stomach cancers and in parallel with BDISH in 152 of these cases. Interestingly, all IHC3+ tumors showed *HER2* gene amplification.

Recommendation 3

Use of FDA approved antibodies is recommended for selection of *HER2*-positive patients in gastric cancer. Besides HercepTest which was applied in the ToGA study another antibody approved by the FDA for breast cancer testing (4B5) appeared to be at least as sensitive, possibly showing even higher inter-laboratory concordance for *HER2* IHC scoring and a closer relationship between *IHC3+* and *HER2* gene amplification.

Given the evolving and provisional status of *HER2* status determination for gastric cancer, we make a final recommendation as follows.

Recommendation 4

Participation in proficiency testing tools such as QUIP (Qualitätssicherungs-Initiative der Deutschen Gesellschaft für Pathologie und des Berufsverbandes Deutscher Pathologen zur diagnostischen Immunhistochemie und Molekularpathologie) at the Dresden Laboratory is strongly recommended since *HER2* testing is quite different in gastric cancer as compared to breast cancer (www.ringversuch.de).

We could show that application of the consented guidelines to a total of 447 diagnostic gastric cancer specimens resulted in a positivity rate of 22.8% which is quite in the range of the ToGA study with 22.1% [25]. Within a subset of 153 tumors, all cases were tested in parallel by IHC and ISH. A complete concordance between both methods could be demonstrated within the *IHC3+* group being all amplified by BDISH. Even so, according to ToGA trial data amplification was not sufficient enough to reliably detect the patients that had a significant benefit from trastuzumab therapy [22]. Although most of amplified *IHC0* and *IHC1+* cases had low level amplification (ratio 2–3) both in ToGA and in our own validation series up to now it is not quite clear whether there is a predictive correlation between therapy response and amplification level in gastric cancer which is obviously not the case in breast cancer [30].

Finally, it turned out that due to heterogeneity of at least some advanced GC both in the mixed-type and in the intestinal-type precise scanning of tumors is of importance particularly if FISH techniques at high magnification ($\times 100$) are used. This might be in favor for light-microscopically based ISH techniques where even small amplified tumor cell foci can readily be recognized (data not shown).

Acknowledgments This validation study was funded by F. Hoffmann-La Roche Ltd (Roche), Basel, Switzerland. The authors thank L. Miller for medical writing support.

Conflict of interest J.R., M.D., G.B., A.W., G.M., M.P.C., F.PL, W.S., H.H.K. were reimbursed for their attendance of the consensus

meeting and for consumables used for testing of the validation set of diagnostic gastric cancer specimens by F. Hoffmann-La Roche Ltd (Roche), Basel, Switzerland.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Simon R, Nocito A, Hübscher T, Bucher C, Torhorst J, Schraml P, Bubendorf L, Mihatsch MM, Moch H, Wilber K, Schötzau A, Kononen J, Sauter G (2001) Patterns of her-2/neu amplification and overexpression in primary and metastatic breast cancer. *J Natl Cancer Inst* 93:1141–1146
2. Vincent-Salomon A, Pierga JY, Couturier J, d'Enghien CD, Nos C, Sigal-Zafrani B, Lae M, Fréneaux P, Diéras V, Thiéry JP, Sastre-Garau X (2007) *HER2* status of bone marrow micrometastasis and their corresponding primary tumours in a pilot study of 27 cases: a possible tool for anti-*HER2* therapy management? *Br J Cancer* 96:654–659
3. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the *HER-2/neu* oncogene. *Science* 235:177–182
4. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A et al (1989) Studies of the *HER-2/neu* proto-oncogene in human breast and ovarian cancer. *Science* 244:707–712
5. Ross JS, Fletcher JA (1998) The *HER-2/neu* oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. *Oncologist* 3:237–252
6. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L (2001) Use of chemotherapy plus a monoclonal antibody against *HER2* for metastatic breast cancer that overexpresses *HER2*. *N Engl J Med* 344:783–792
7. Marty M, Cognetti F, Maraninchi D, Snyder R, Mauriac L, Tubiana-Hulin M, Chan S, Grimes D, Antón A, Lluca A, Kennedy J, O'Byrne K, Conte P, Green M, Ward C, Mayne K, Extra JM (2005) Randomized phase II trial of the efficacy and safety of trastuzumab combined with docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer administered as first-line treatment: The M77001 study group. *J Clin Oncol* 23:4265–4274
8. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA, Swain SM, Pisansky TM, Fehrenbacher L, Kutteh LA, Vogel VG, Visscher DW, Yothers G, Jenkins RB, Brown AM, Dakhil SR, Mamounas EP, Lingle WL, Klein PM, Ingle JN, Wolmark N (2005) Trastuzumab plus adjuvant chemotherapy for operable *HER2*-positive breast cancer. *N Engl J Med* 353:1673–1684
9. Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M, Goldhirsch A, Untch M, Mariani G, Baselga J, Kaufmann M, Cameron D, Bell R, Bergh J, Coleman R, Wardley A, Harbeck N, Lopez RI, Mallmann P, Gelmon K, Wilcken N, Wist E, Sánchez Rovira P, Piccart-Gebhart MJ; HERA study team (2007) 2-year follow-up of trastuzumab after adjuvant chemotherapy in *HER2*-positive breast cancer: a randomised controlled trial. *Lancet* 369:29–36

10. Arnould L, Arveux P, Couturier J, Gelly-Marty M, Loustalot C, Ettore F, Sagan C, Antoine M, Penault-Llorca F, Vasseur B, Fumoleau P, Coudert BP (2007) Pathological complete response to trastuzumab-based neoadjuvant therapy is related to the level of HER-2 amplification. *Clin Cancer Res* 13:6404–6409
11. Hofmann M, Stoss O, Shi D, Büttner R, van de Vijver M, Kim W, Ochiai A, Rüschoff J, Henkel T (2008) Assessment of a HER2 scoring system for gastric cancer: results from a validation study. *Histopathology* 52:797–805
12. 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 (2006) HER-2/neu amplification is an independent prognostic factor in gastric cancer. *Dig Dis Sci* 51:1371–1379
13. Yano T, Doi T, Ohtsu A, Boku N, Hashizume K, Nakanishi M, Ochiai A (2006) Comparison of HER2 gene amplification assessed by fluorescence in situ hybridization and HER2 protein expression assessed by immunohistochemistry in gastric cancer. *Oncol Rep* 15:65–71
14. Zhang XL, Yang YS, Xu DP, Qu JH, Guo MZ, Gong Y, Huang J (2009) Comparative study on overexpression of her2/neu and her3 in gastric cancer. *World J Surg* 33:2112–2118
15. Uchino S, Tsuda H, Maruyama K, Kinoshita T, Sasako M, Saito T, Kobayashi M, Hirohashi S (1993) Overexpression of c-erbB-2 protein in gastric cancer. Its correlation with long-term survival of patients. *Cancer* 72:3179–3184
16. Nakajima M, Sawada H, Yamada Y, Watanabe A, Tatsumi M, Yamashita J, Matsuda M, Sakaguchi T, Hirao T, Nakano H (1999) The prognostic significance of amplification and overexpression of c-met and c-erb B-2 in human gastric carcinomas. *Cancer* 85:1894–1902
17. Allgayer H, Babic R, Gruetzner KU, Tarabichi A, Schildberg FW, Heiss MM (2000) 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* 18:2201–2209
18. García I, Vizoso F, Martín A, Sanz L, Abdel-Lah O, Raigoso P, García-Muñiz JL (2003) Clinical significance of the epidermal growth factor receptor and HER2 receptor in resectable gastric cancer. *Ann Surg Oncol* 10:234–241
19. Tanner M, Hollmén M, Junttila TT, Kapanen AI, Tommola S, Soini Y, García-Muñiz JL (2005) Amplification of HER-2 in gastric carcinoma: association with topoisomerase IIalpha gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. *Ann Oncol* 16:273–278
20. Matsui Y, Inomata M, Tojigamori M, Sonoda K, Shiraishi N, Kitano S (2005) Suppression of tumor growth in human gastric cancer with HER2 overexpression by an anti-HER2 antibody in a murine model. *Int J Oncol* 27:681–685
21. Fujimoto-Ouchi K, Sekiguchi F, Yasuno H, Moriya Y, Mori K, Tanaka Y (2007) Antitumor activity of trastuzumab in combination with chemotherapy in human gastric cancer xenograft models. *Cancer Chemother Pharmacol* 59:795–805
22. Van Cutsem E, Kang Y, Chung H, Sawaki A, Lordick F, Hill J, Lehle M, Feyereislova A, Bang Y (2009) Efficacy results from the ToGA trial: A phase III study of trastuzumab added to standard chemotherapy (CT) in first-line human epidermal growth factor receptor 2 (HER2)-positive advanced gastric cancer (GC). *J Clin Oncol* 27:18 s (suppl; abstr LBA4509)
23. Bang Y, Van Cutsem C, Feyereislova A, Chung H, Shen L, Sawaki S, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, Hill J, Lehle M, Rüschoff J, Kang Y-K (2010) Phase III, randomised clinical trial to compare trastuzumab in combination with chemotherapy versus chemotherapy alone for the treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer. *Lancet*, submitted
24. EMEA, European Medicines Agency (2009): Opinion www.emea.europa.eu/pdfs/human/opinion/Herceptin_82246709en.pdf
25. Bang Y, H. Chung H, Xu J, Lordick F, Sawaki A, Al-Sakaff N, Lipatov O, See C, Rueschoff J, Van Cutsem E (2009) Pathological features of advanced gastric cancer (GC): relationship to human epidermal growth factor receptor 2 (HER2) positivity in the global screening programme of the ToGA trial. *J Clin Oncol* 27:15s (suppl; abstr 4556)
26. Dietel M, Ellis IO, Höfler H, Kreipe H, Moch H, Dankof A, Kölbl K, Kristiansen G (2007) 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* 451:19–25
27. Barros-Silva JD, Leitão D, Afonso L, Vieira J, Dinis-Ribeiro M, Fragoso M, Bento MJ, Santos L, Ferreira P, Rêgo S, Brandão C, Carneiro F, Lopes C, Schmitt F, Teixeira MR (2009) Association of ERBB2 gene status with histopathological parameters and disease-specific survival in gastric carcinoma patients. *Br J Cancer* 100:487–493
28. Tapia C, Glatz K, Novotny H, Lugli A, Horcic M, Seemayer CA, Tornillo L, Terracciano L, Spichtin H, Mirlacher M, Simon R, Sauter G (2007) Close association between HER-2 amplification and overexpression in human tumors of non-breast origin. *Mod Pathol* 20:192–198
29. Marx AH, Tharun L, Muth J, Dancau AM, Simon R, Yekebas E, Kaifi JT, Mirlacher M, Brümmendorf TH, Bokemeyer C, Izbicki JR, Sauter G (2009) HER-2 amplification is highly homogenous in gastric cancer. *Hum Pathol* 40:769–777
30. Dowsett M, Procter M, McCaskill-Stevens W, de Azambuja E, Dafni U, Rueschoff J, Jordan B, Dolci S, Abramovitz M, Stoss O, Viale G, Gelber RD, Piccart-Gebhart M, Leyland-Jones B (2009) Disease-free survival according to degree of HER2 amplification for patients treated with adjuvant chemotherapy with or without 1 year of trastuzumab: the HERA trial. *J Clin Oncol* 27:2962–2969
31. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 25:118–145
32. Rüschoff J, Nagelmeier I, Baretton G, Dietel M, Höfler H, Schildhaus HU, Büttner R, Schlake W, Stoss O, Kreipe HH (2010) HER2 testing in gastric cancer. What is different in comparison to breast cancer? *Pathologie* 31:208–217 [German]
33. Conger AJ (1980) Integration and generalization of Kappa for multiple raters. *Psychol Bull* 88:322–328