Digital image analysis of HER2 immunostained gastric and gastroesophageal junction adenocarcinomas

Sofie Lunden Nielsen, MS, Søren Nielsen, BSC, and Mogens Vyberg, MD
Institute of Pathology, Department of Clinical Medicine, Aalborg University Hospital

Abstract

Background: Manual assessment of HER2 protein expression in gastric and gastroesophageal junction (GEJ) adenocarcinoma is prone to inter-observer variability and hampered by tumor heterogeneity and different scoring criteria. This study aimed to evaluate the accuracy of digital image analysis (DIA) as a more objective and precise method for the assessment of HER2 protein expression.

Methods: Hundred and ten gastric and GEJ adenocarcinomas were included applying a tissue micro array (TMA) approach with three cores per case. Two immunohistochemistry (IHC) assays, PATHWAY® and HercepTest™, and fluorescence in situ hybridization (FISH) were performed for all cases. Manual interpretation of IHC slides followed guidelines for both resection and biopsy specimens. The HER2 CONNECT™ DIA software as designed for breast carcinoma was applied. Connectivity, calculated by the software, was converted to the standard IHC scores (negative, equivocal, positive) applying predetermined cut-off values for breast carcinoma as well as novel cut-off values. Cases with excessive cytoplasmic and nuclear staining as well as HER2 amplified IHC negative cases were excluded from HER2 CONNECT™ analysis.

Results: Manual scoring of IHC slides, using criteria for biopsies, achieved the most optimal combination of specificity and sensitivity using FISH as the reference. Applying HER2 CONNECT™ with established connectivity cut-off values designed for breast carcinoma resulted in 72.7% sensitivity and 100% specificity for the identification of HER2 amplified cases. The sensitivity was increased to 100% when the new cut-off values were applied, while the specificity remained 100%. With the new cut-off values, a statistically significant reduction of IHC equivocal cases (50.0% for

PATHWAY® and 36.4% for HercepTest™) was observed. Three cores with HER2 protein overexpression and amplification were classified as negative by HER2 CONNECT™. However, the other cores from the specific cases ensured an accurate classification.

Conclusion: HER2 CONNECT™ seems to be an effective tool for assessment of HER2 protein expression and gene amplification in gastric and GEJ adenocarcinoma.

Resumé

Overekspression af human epidermal growth factor receptor 2 (HER2) forekommer i omtrent 18% af alle adenokarcinomer i ventriklen og den gastroesofageale overgang (GEJ). Ved HER2 overekspression og samtidig metastatisk sygdom kan patienten tilbydes behandling med trastuzumab (Herceptin). HER2 protein ekspressionen i pågældende tumorer skal vurderes for at finde patienter, der er kandidater til behandlingen. Indledningsvist udføres immunhistokemi (IHC). Bedømmelse af IHC snittene følger fastsatte kriterier, hvorfor det kan afgøres, om tumoren er negativ eller positiv for HER2 overekspression. Endvidere vil nogle tumorer blive klassificeret som borderline (equivocal). Disse cases er tvivlstilfælde, hvor det er uvist om patienten vil have gavn af behandling med trastuzumab. Her er det nødvendigt at foretage gentest med en in situ hybdridiserings (ISH) metode, hvor der undersøges for HER2 genamplifikation.

Vurdering af HER2 ekspressionen besværliggøres af heterogent tumorvæv, dårlig inter-observatør overensstemmelse og forskellige scoringskriterier, afhængigt af, om det er biopsier og resektater, der undersøges. Endvidere er fortolkning af både IHC og ISH præget af en vis subjektivitet. En mere objektiv og reproducerbar analysemetode af IHC er dermed eftertragtet. Anvendelse af digital billedanalyse (DIA) til vurdering af HER2 ekspression er blevet godkendt til diagnostisk brug inden for mammakarcinomer, hvor HER2 CONNECT™ (DIA software) er dokumenteret som et anvendeligt diagnostisk værktøj med høj analytisk præcision og medfører samtidig en reduktion af antallet af borderline cases. Dette studie havde til formål at evaluere præcisionen af HER2 CONNECT™ ved analyse af HER ekspression i adenokarcinomer i ventrikel og GEJ.

110 ventrikel- og GEJ-adenokarcinomer blev inkluderet i studiet. Studiet anvendte et tissue micro array (TMA) set up, hvor hver tumor blev repræsenteret af tre cores. To IHC protokoller for HER2 protein blev anvendt i form af PATHWAY® (Ventana) og HercepTest™ (Dako). IHC snittene blev fortolket ved brug af scoringskriterier for både resektater og biopsier, da det endnu ikke er blevet undersøgt, hvilke af disse kriterier, der skal anvendes ved TMA'er. Fluorescens in situ hybridisering

(FISH) for HER2 genet blev udført på alle cases. HER2 CONNECT™ analyserede alle cases med udregning af en connectivity værdi for hver case. Ud fra bestemte cut-off værdier, blev connectivity omdannet til DIA scores sammenlignelige med standard IHC scores (negativ, equivocal og positiv). Cut-off værdier fastlagt for mammakarcinomer blev først appliceret og efterfølgende blev nye cut-off værdier fundet og anvendt ved sammenligning af connectivity data med manuelt aflæst IHC og FISH resultat. Cases med kerne- og cytoplasmafarvning samt IHC negative HER2 genamplificerede cases blev ekskluderet fra HER2 CONNECT™ analysen.

IHC blev fortolket ved brug af kriterier for både resektater og biopsier, hvorefter resultater fra de to sæt blev sammenlignet. Her blev det fundet, at kun brug af kriterier for biopsier resulterede i 100% sensitivitet. Dette gjorde sig gældende for både PATHWAY® og HercepTest™. Specificiteten var 100% ved begge IHC protokoller og begge scoringsmetoder. Dermed blev kun resultater fra scoring ved brug af guidelines for biopsier anvendt i den efterfølgende databehandling.

Brug af cut-off værdier bestemt for mammakarcinomer resulterede i 72,7% sensitivitet og 100% specificitet for HER2 CONNECT™ ved begge IHC protokoller. Anvendelse af de nye cut-off værdier resulterede i 100% sensitivitet og specificitet for HER2 CONNECT™ ved begge IHC protokoller. Desuden blev der fundet en statistisk signifikant ($p < 0,05$) reduktion af antallet af equivocal cases svarende til 50,0% for PATHWAY® og 36,4% HercepTest™. Tre cores med HER2 protein overekspression og genamplifikation blev klassificeret som negative af HER2 CONNECT™. De andre cores fra pågældende cases sikrede dog en korrekt klassifikation.

Det konkluderes, at HER2 CONNECT™ synes at være et præcist og effektivt diagnostisk værktøj til vurdering af HER2 protein ekspression i ventrikel- og GEJ-adenokarcinomer. Studiets resultater skal valideres i yderligere studier med inddragelse af flere patologiafdelinger.

Introduction

Gastric and esophageal cancer account for the third and sixth most common cause, respectively, of cancer death worldwide.[1] These cancers show poor prognosis with a 5-year survival rate of 28 and 18% respectively, disregarding stage of disease at diagnosis.[2] The only potentially curable treatment is surgery. However, at the time of diagnosis most patients present with inoperable disease (except in countries with screening programs).[3-5] Inoperable patients and patients with recurrent and metastatic cancer are submitted to palliative treatment. Approximately 18% of gastric and gastroesophageal junction (GEJ) adenocarcinomas exhibit the tyrosine kinase human epidermal growth factor receptor 2 (HER2) overexpression[6], enabling treatment with trastuzumab in combination with chemotherapy. Trastuzumab is a monoclonal antibody targeting HER2. It prevents activation of the tyrosine kinase by multiple possible mechanisms including antibody-dependent cytotoxicity, prevention of receptor dimerization, blocking cleavage of the extracellular domain and promoting endocytic destruction of the receptor.[7] The US Food and Drug Administration (FDA) and European Medicines Agency (EMA) have granted their approval of trastuzumab treatment for gastric and GEJ cancers based on the ToGA trial[8,9], which demonstrated a statistically significant increase in median overall survival in patients treated with trastuzumab in combination with chemotherapy compared to chemotherapy alone.[10]

HER2 expression is primarily assessed by immunohistochemistry (IHC), using defined and validated scoring criteria in order to identify patients eligible for trastuzumab treatment. The level of HER2 protein expression is assessed semi-quantitatively by evaluation of intensity and percentage of stained tumor cells. Cases are scored on a scale of 0–3+ where scores of 0 and 1+ are categorized as negative, 2+ as equivocal, and 3+ as positive.[11,12] However, this method is complicated by several pitfalls, including

technical issues (different antibodies, IHC protocols and stainer platforms giving varying staining reactions)[12-14], inter-observer variability [12,13] and tumor heterogeneity [11,12]. Equivocal cases need further analysis by fluorescence in situ hybridization (FISH) or bright field in situ hybridization (BRISH) methods. These methods are less cost-effective and require more expertise to conduct than IHC assays. According to EMA, treatment with trastuzumab should only be administered for patients/cases with positive IHC result or IHC equivocal cases with confirmed HER2 gene amplification, whereas HER2 IHC negative and gene amplified cases are not considered eligible for trastuzumab[8].

Manual interpretation of IHC assays is associated with a certain amount of subjectivity, which contributes to inter-observer variability. A more objective analysis method for evaluating HER2 protein expression in gastric and GEJ adenocarcinomas can potentially be achieved by application of digital image analysis (DIA). This method is recommended for breast cancer by American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) as an effective tool for interpretation.[15] The DIA company Visiopharm has in collaboration with the Institute of Pathology, Aalborg University Hospital, developed and commercialized a software, HER2 CONNECT™ enabling an accurate and effective analysis of HER2 IHC status in breast cancer cases.[16] Similar software to assess HER2 IHC status in gastric and GEJ adenocarcinomas has not yet been validated and launched. Few previous studies have evaluated the use of DIA for HER2 expression in gastric and GEJ adenocarcinomas. These studies applied algorithms validated for breast cancer and did not exhibit optimal concordance rates with manual IHC interpretation and FISH.[17,18] One study applied a software algorithm designed for gastric cancer. However, concordance between DIA and manual IHC and FISH result was not included in the paper.[19]

The aim of the present study was to evaluate the accuracy of HER2 CONNECT™ as analysis tool for interpretation of HER2 protein expression in gastric and GEJ adenocarcinomas based on HercepTest™ and PATHWAY® with FISH as reference.

Materials and Methods

The material consisted of 110 consecutive resection specimens (RS) from gastric and GEJ adenocarcinomas with sufficient amounts of tumor tissue for examination, collected retrospectively from the archives of Institute of Pathology, Aalborg University Hospital, during 2002-2015. No inclusion/exclusion criteria regarding neoadjuvant chemotherapy were employed.

All specimens were subjected to standard processing methods including fixation in 10% neutral buffered formalin for 24-72 hours.

For the present study, 11 TMA blocks were constructed (TMA master, 3DHISTECH) as follows: From each of the 110 specimens, three tumor-containing regions were identified on hematoxylin-eosin (HE) stained full slides and punched out of the paraffin blocks with a 2.0 mm needle. Each TMA further included two tissue cores of breast ductal carcinomas as run controls, one characterized as IHC HER2 equivocal and one IHC HER2 positive. Three liver tissue cores were included to ensure navigation during the interpretation. Consecutive 4 µm sections were cut and mounted onto coated slides (FLEX IHC slides K8020, Dako). The sections were dried overnight at room temperature and then stored at -20°C until staining. Sections from all TMA blocks were stained simultaneously using the same reagents (lot numbers etc.) and protocol settings to minimize inter-run variations.

The following assays were applied:

A. HER2 IHC - HercepTest™
B. HER2 IHC - PATHWAY®
C. HER2 FISH - Zyto*Light*®

A. IHC; HercepTest™ (Dako, SK001)

Slides were stained according to the recommendations described in the package insert of the kit (Dako, SK001), and in brief processed as follows: slides were dried at 60°C

for one hour before deparaffinization in Tissue-Clear (Sakura) and hydrated through alcohol to distilled water. The slides were then submitted to heat-induced epitope retrieval (HIER) for 40 minutes at 97°C in PT-link (Dako). After cooling down for 20 minutes, the slides were placed in the Autostainer Link 48 (Dako) and subsequently the immunohistochemical procedure was performed. After blocking for endogenous peroxidase for 5 minutes and wash in buffer, the slides were incubated with the primary antibody (rabbit polyclonal, Dako SK001) at room temperature for 30 minutes. Following wash in buffer, the visualization complex (horseradish peroxidase (HRP)-labeled polymer, Dako, SK001) was applied for 30 minutes. After a wash in the buffer the slides were finally developed with 3,3'Diaminobenzidine tetrahydrocholoride (DAB) (Dako, SK001) and counterstained with Mayers hematoxylin (S3301, Dako).

B. IHC; PATHWAY® (Ventana, 790-2991)

According to the manufacturer's recommendations, the slides were dried at 60°C for one hour and placed in the BenchMark Ultra instrument (Ventana). The slides were deparaffinized on-board and submitted to HIER in Cell Conditioning 1 for 32 minutes at 95°C. Following endogenous peroxidase blocking, the primary antibody (rabbit monoclonal clone 4B5, 760-2991) was applied for 20 minutes at 36°C. After a wash in buffer the visualization complex, UltraView DAB (HRP-labeled multimer, Ventana, 760-500) was then applied and after a wash in the buffer, the slides were finally developed with DAB (Ventana, 760-500) and counterstained with hematoxylin II (Ventana, 790-2208).

C. FISH; Zyto*Light*® (Zytovision, Z-2015)

The sections were dried at 60°C overnight before deparaffinization in xylene and hydration through alcohol to distilled water. The specimens were heated in a pre-treatment solution (Dako, K5799) in a domestic microwave oven (Blomberg) for 10 minutes. Hereafter, the slides were submitted to

proteolytic digestion using pepsin (Dako, K5799) at room temperature for 10 minutes. Denaturation for four minutes at 90°C and hybridization for 16 hours at 37°C was performed in a hybridizer (Dako, S2450). The probes for hybridization were based on a dual-probe mix (Zytovision, Z-2015) containing a mixture of an orange fluorochrome direct labeled probe specific for the alpha satellite centromeric region of chromosome 17 and a green fluorochrome direct labeled probe specific for the chromosomal region 17q12-q21.1 harboring the HER2 gene. After a stringent wash at 65°C for 10 minutes, the slides were mounted with a fluorescence mounting medium containing 4',6-Diamidino-2-phenylindol dihydrocholoride and cover-slipped. The slides were stored at 2°C to 8°C in the dark until evaluation, which was terminated within two weeks. The results were interpreted using a fluorescence microscope (Leica DMRXA).

Interpretation of IHC assays

Interpretation of the IHC assays followed the validated guidelines for gastric and GEJ cancer (table 1).[11,12] The criteria differ between RS and biopsy specimens (BS). Previous studies with a TMA set-up have used both scoring criteria for RS[12,19-22]. Therefore, this study applied both criteria to evaluate correlation with FISH.

Table 1: Scoring guidelines for interpretation of IHC

Resection specimen staining pattern	Biopsy specimen staining pattern	Score	Classification
No reactivity or membranous reactivity in < 10% of cells	No reactivity or membranous staining in any tumor cell	0	Negative
Faint/barely perceptible membranous reactivity in ≥ 10% of tumor cells (at 40x)	Tumor cluster of ≥ 5 cells with faint/barely perceptible membranous reactivity (at 40x)	1+	Negative
Weak/moderate complete or basolateral membranous reactivity in ≥ 10% of tumor cells (at 10/20x)	Tumor cluster of ≥ 5 cells with weak/moderate complete or basolateral membranous reactivity (at 10/20x)	2+	Equivocal
Moderate/strong complete or basolateral membranous reactivity in ≥ 10% of tumor cells (at 2.5/5x)	Tumor cluster of ≥ 5 cells with moderate/strong complete or basolateral membranous reactivity (at 2.5/5x)	3+	Positive

Two observers scored the IHC slides. Discrepant cases were reevaluated to achieve consensus. The core with the highest score was considered the final result for the case.

Intra-tumoral heterogeneity was evaluated. A case was defined as heterogeneous when it consisted of 1) both negative and equivocal/ positives scores and 2) both equivocal and positive scores.

Furthermore, non-specific staining i.e. cytoplasmic and nuclear staining of tumor cells was assessed. A case was classified as equivocal if ≥ 10% of the tumor cells (RS criteria) or a cluster of ≥ 5 tumor cells (BS criteria) could not be assessed with certainty due to this aberrant staining pattern.

Interpretation of FISH

HER2 gene amplification was classified as negative if HER2/CEN17 ratio was < 2 and positive if HER2/CEN17 ratio being ≥ 2. For each core, 20 non-overlapping representative nuclei were counted. FISH interpretation was conducted by an experienced biomedical laboratory scientist. FISH was evaluated in "hot spots" and if the result was discordant with IHC, the case was recounted

(enumeration of 20 additional nuclei) using IHC to identify regions of interest.

Digital image analysis

TMA slides were scanned with NanoZoomer HT 1.0 (Hamamatsu) at 40x magnification to obtain digital images available for automated image analysis.

The TMA Workflow module of the HER2 CONNECT™ software was used to create a TMA template, which fitted the design of the TMAs included in this study. An individual image for each core was automatically generated by fitting the template for each TMA.

The HER2 CONNECT™ software algorithm applies following steps to determine a connectivity value, which is later translated to a score equivalent to standard IHC scores: [16,23]

1. Pre-processing: identification of brown pixels, which are part of linear structures.
2. Segmentation: selection of relevant pixels based on the intensity of the brown color and dimensions of linearity.
3. Post-processing: skeletonization of membrane segments to a thickness of one pixel, exclusion of small membrane fragments according to a specified cut-off value, and merging of other membrane fragments to obtain perfect connection.

The size distribution of the membrane segments, which are present after the steps mentioned above, determines the connectivity.

Connectivity values were converted to negative (0/1+), equivocal (2+) and positive (3+) DIA scores primarily according to validated cut-off values determined for breast carcinoma specimens (table 2)[24]. Secondary, connectivity values were compared to IHC and FISH results to determine new cut-off values, specifically adjusted for gastric and GEJ adenocarcinoma specimens.

Table 2: Cut-off values for breast carcinomas

Connectivity	Equivalent IHC score
≤ 0.40	0/1+
> 0.40 - ≤ 0.64	2+
> 0.64	3+

A new classification, called HER2 score, combining the IHC and FISH result was given for each core and case to simultaneously compare connectivity to both IHC and FISH results (table 3).

Table 3: HER2 score combining IHC and FISH score

HER2 Score	IHC score	FISH score
0	0	neg
1.0	1+	neg
1.5	2+	neg
2.5	2+	pos
3.0	3+	pos

The highest connectivity value and HER2/CEN17 ratio was selected from the three cores to represent the case.

Statistical analysis

IBM SPSS Statistics 23.0 was used as statistical software. Firstly, results from IHC interpretation using criteria for RS and BS were compared to FISH. The method providing highest analytical sensitivity and specificity was selected. Analytical sensitivity, specificity and accuracy were calculated for both IHC assays interpreted manually and by HER2 CONNECT™ using FISH as reference. McNemar's test was employed to evaluate the difference between manual and automatic IHC assessment. Cohen's kappa was calculated to analyze agreement between PATHWAY® and HercepTest™ in relation to HER2 CONNECT™ results.

Results

IHC and FISH results were available in 104 of the 110 cases included, for both PATHWAY® and HercepTest™. One case was excluded from all assays as no tumor cells could be identified. Five cases were not evaluable by FISH due to technical issues. Additional two and three cores were excluded for the PATHWAY® and HercepTest™ assays, respectively, because of poor tumor tissue quality. The affected cases were thus represented by only one or two cores.

Comparison of manual IHC interpretation to FISH results

Results from IHC interpretation, applying RS and BS scoring criteria, were compared to FISH (table 4).

Two cases were scored as 0 for PATHWAY® and HercepTest™ in IHC interpretation (when criteria for both RS and BS were applied), but were both amplified by FISH. The two cases had low-level amplification (HER2/CEN17 ratio of 2.1-2.2) and showed heterogeneous amplification as only one and two cores, respectively, were amplified by FISH. Interpretation based on criteria for RS lead to six false negative cases for PATHWAY® and three for HercepTest™.

Table 4: Comparison of manual interpretation of IHC to FISH results

PATHWAY® Resection	FISH		Total	HercepTest™ Resection	FISH		Total
	neg	pos			neg	pos	
0	60	3	63	**0**	65	3	68
1+	15	3	18	**1+**	16	0	16
2+	14	3	17	**2+**	8	5	13
3+	0	6	6	**3+**	0	7	7
Total	89	15	104	**Total**	89	15	104
PATHWAY® Biopsy	**FISH**		**Total**	**HercepTest™ Biopsy**	**FISH**		**Total**
	neg	**pos**			**neg**	**pos**	
0	45	2	47	**0**	50	2	52
1+	19	0	19	**1+**	20	0	20
2+	25	2	27	**2+**	19	3	22
3+	0	11	11	**3+**	0	10	10
Total	89	15	104	**Total**	89	15	104

Sensitivity, specificity and accuracy were calculated for all assays (table 5), excluding the 2+ cases. Scoring TMAs analogous to biopsies provided the highest analytical sensitivity of 84.6% for PATHWAY® and 83.3% for HercepTest™. Specificity was 100% for all scoring methods and IHC assays. Analytical accuracy increased with use of criteria for BS compared to criteria for RS; 97.4% and 97.6% was obtained by criteria for BS compared to 93.1% and 96.7% for

PATHWAY® and HercepTest™, respectively, using criteria for RS.

Table 5: Sensitivity, specificity and accuracy for manual interpretation of IHC compared to FISH results.

	Sensitivity	Specificity	Accuracy
PATHWAY® Resection	50.0%	100%	93.1%
HercepTest™ Resection	70.0%	100%	96.7%
PATHWAY® Biopsy	84.6%	100%	97.4%
HercepTest™ Biopsy	83.3%	100%	97.6%

The most optimal combination of sensitivity and specificity was thus attained when TMAs were scored analogous to biopsies. Therefore, only results from using criteria for BS were included in the following data analysis.

The sensitivity increased to 100% for both PATHWAY® and HercepTest™ when excluding the two cases, which were IHC negative but amplified by FISH. According to EMA, HER2 protein overexpression (2+/3+) must be identified in gene amplified cases in order to apply trastuzumab as treatment[8].

Prevalence of HER2 overexpression and amplification

For PATHWAY®, 66 (63.4%) of the cases were scored as negative, 27 (26.0%) as equivocal and 11 (10.6%) as positive. For HercepTest™, 72 (69.2%) of the cases were scored as negative, 22 (21.2%) as equivocal and 10 (9.6%) as positive (see figure 1 for examples on the different IHC scores).

The prevalence of HER2 gene amplification assessed by FISH was 14.4% (15 cases). All 3+ cases for PATHWAY® and HercepTest™ were amplified by FISH. Two of 27 equivocal cases for PATHWAY® and three of 22 cases for HercepTest™ were positive for HER2 gene amplification. As mentioned, two cases scored as 0 for both PATHWAY® and HercepTest™ had low level HER2 gene amplification.

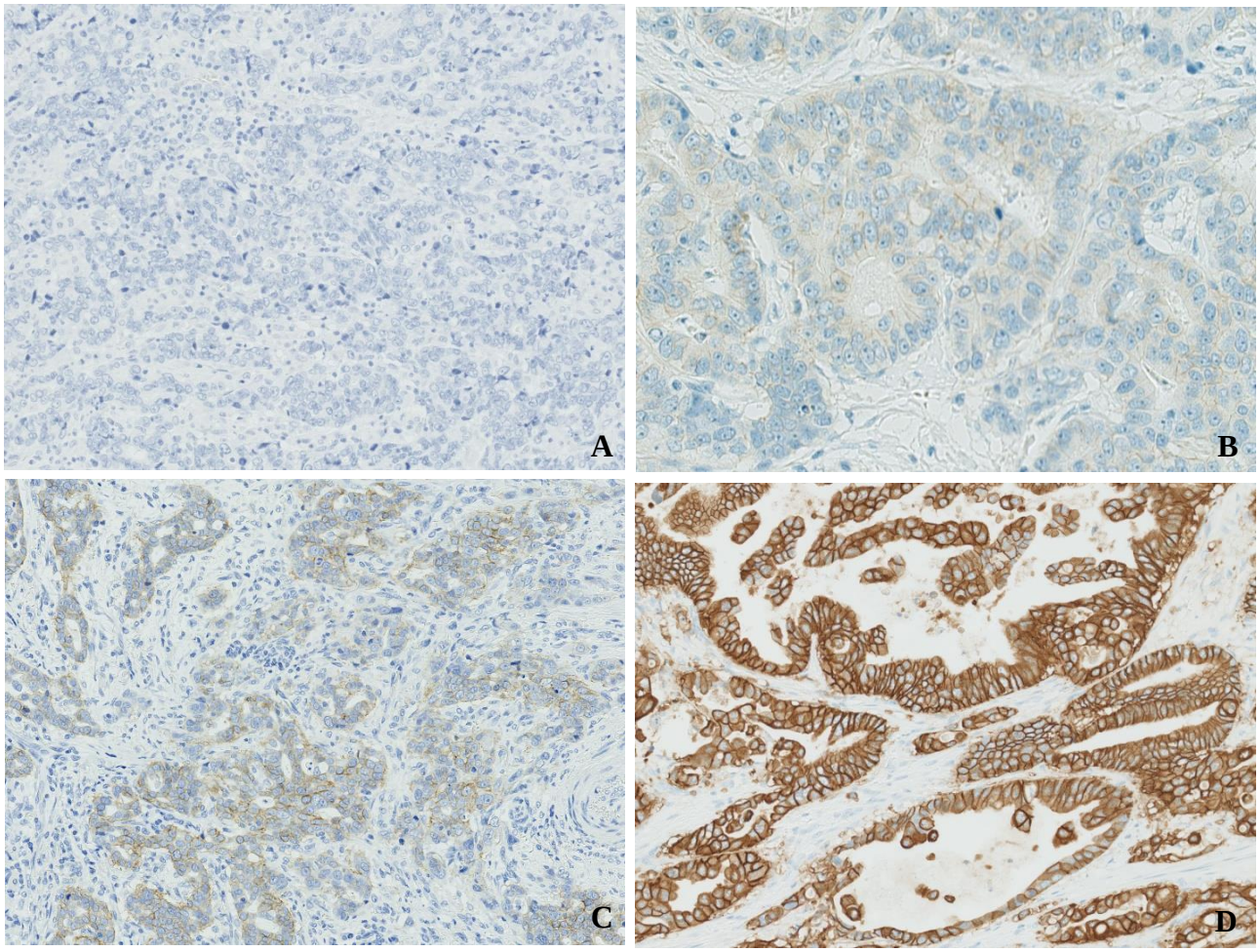

Figure 1: Examples of different HER2 IHC scores (BS criteria) **A** IHC score 0, **B** IHC score 1+, **C** IHC score 2+, **D** IHC score 3+ (**A-D** 20x magnification)

Non-specific staining

The PATHWAY® assay, based on the primary antibody clone 4B5, occasionally provided an aberrant cytoplasmic and nuclear staining. The aberrant staining was observed in both tumor cells and in normal or dysplastic epithelial cells (see figure 2). For this reason, seven cases (equal to 15 cores) were classified as equivocal for the PATHWAY® assay. All seven cases were scored as IHC negative by HercepTest™.

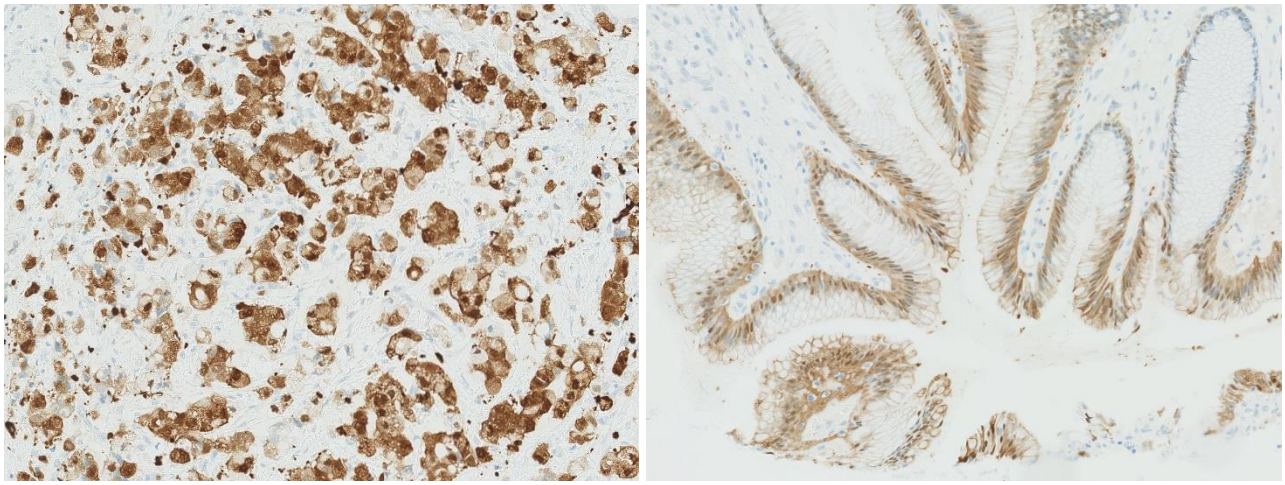

Figure 2: Non-specific staining. **A** PATHWAY®, cytoplasmic and nuclear staining of tumor cells, retesting by FISH is necessary. **B** PATHWAY®, cytoplasmic and nuclear staining of normal epithelial cells. **A-B** 20x magnification

IHC – tumor heterogeneity

Heterogeneous cases, i.e. cases with different IHC scores in the three cores was noted (see figure 3A-3C).

For PATHWAY® and HercepTest™, 18 and 16 cases, respectively, were classified as heterogeneous. This equals to respectively 16.5% and 14.7% of all IHC cases.

Different intensity of membranous staining within the same core was also observed (see figure 3D).

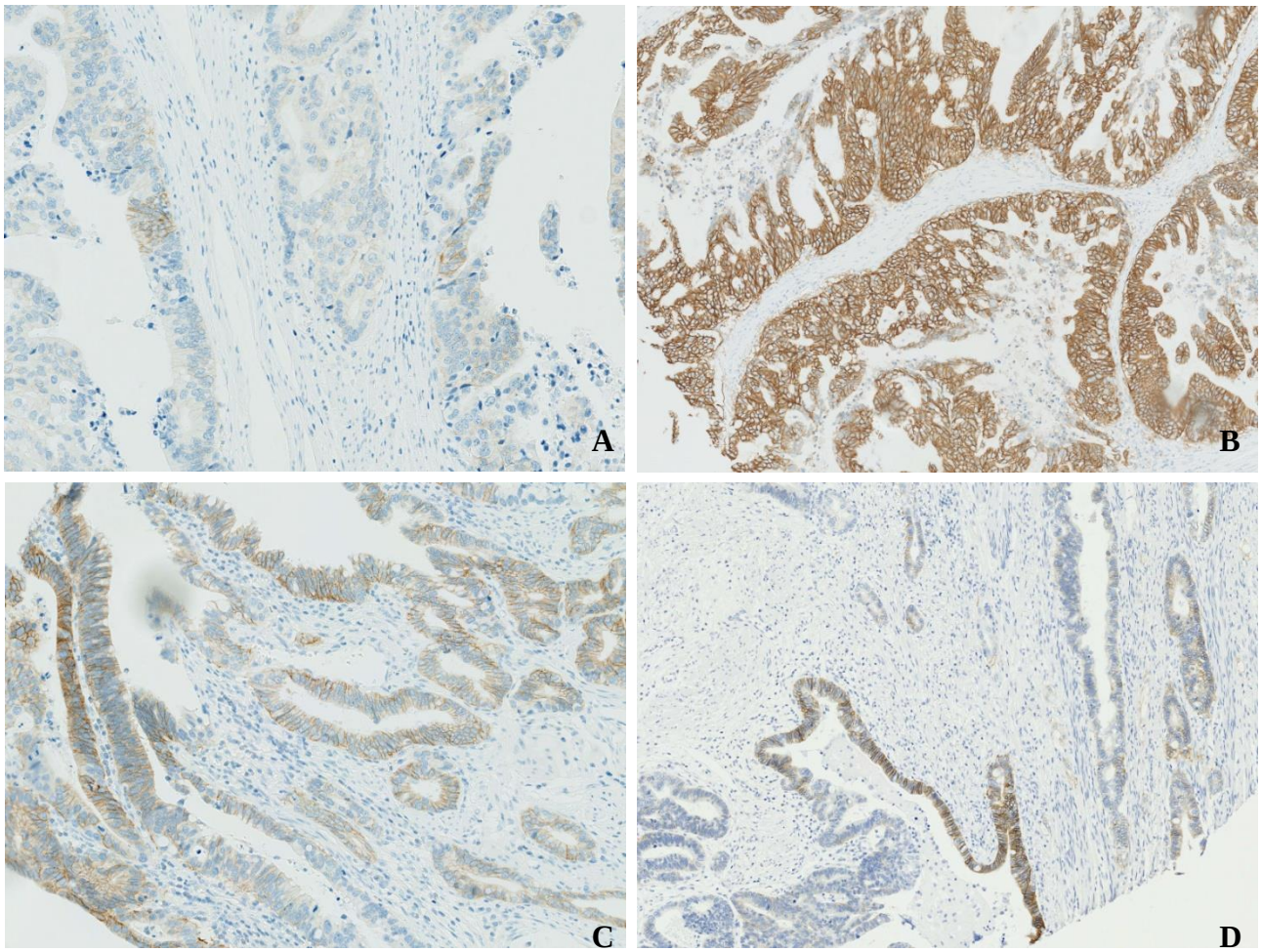

Figure 3: Heterogeneous cases. **A-C,** same case with different IHC scores (BS criteria) in the three individual cores, PATHWAY®. **A** scored as 2+ (only a small cluster of 2+ stained tumor cells), 20x **B** scored as 3+, 10x **C** scored as 3+ (only small clusters of 3+ staining, mostly 2+ staining), 20x **D** HercepTest™, one core with IHC score ranging from 0-3+, 10x

HER2 CONNECT™

All cases and cores with manual IHC negative result and amplification by FISH were excluded from HER2 CONNECT™ analysis, as no membranous staining was available for analysis. For this reason, two cases (equal to three cores) were excluded from both IHC assays, and additional three cores for PATHWAY® and one core for HercepTest™ were also excluded. However, the other cores from the specific cases classified the final result for the case as equivocal or positive in manual IHC interpretation.

Cases with excessive cytoplasmic and nuclear staining requiring retest by FISH were not eligible for HER2 CONNECT™ analysis. The software was not able to distinguish the non-

specific staining from the specific membranous staining. 15 cores equaling 7 cases, for PATHWAY®, were consequently excluded from HER2 CONNECT™ analysis.

Cut-off values, originally determined for breast carcinoma samples (table 2), were applied to convert connectivity to DIA scores in line with the manually applied IHC scores. The DIA scores were compared to FISH for all cases (table 6).

Table 6: Comparison of DIA scores to FISH applying cut-off values designed for breast carcinoma (cases)

		FISH		Total
		neg	pos	
PATHWAY® DIA Score	**0/+1**	82	3	85
	2+	0	2	2
	3+	0	8	8
Total		82	13	95
		neg	**pos**	**Total**
HercepTest™ DIA Score	**0/1+**	89	3	92
	2+	0	2	2
	3+	0	8	8
Total		89	13	102

Three amplified cases were classified as negative by HER2 CONNECT™ for both IHC assays, when cut-offs for breast carcinoma were used. This resulted in a sensitivity of 72.7% and a specificity of 100% for both PATHWAY® and HercepTest™. Accuracy was 96.8% for PATHWAY® and 97.0% for HercepTest™.

A HER2 score as described in table 3, combining manual IHC and FISH results, was given for each core and case (for distribution of HER2 scores, see table 7). Hereafter, the connectivity values were compared to HER2 scores to define the intervals for the new cut offs (see figure 4).

Table 7: Distribution of HER2 scores for all cores

HER2 score	PATHWAY® (%)	HercepTest™ (%)
0	174 (60.2)	184 (60.3)
1.0	41 (14.2)	58 (19.0)
1.5	45 (15.6)	32 (10.5)
2.5	7 (2.4)	10 (3.3)
3.0	22 (7.6)	21 (6.9)
Total	289	305

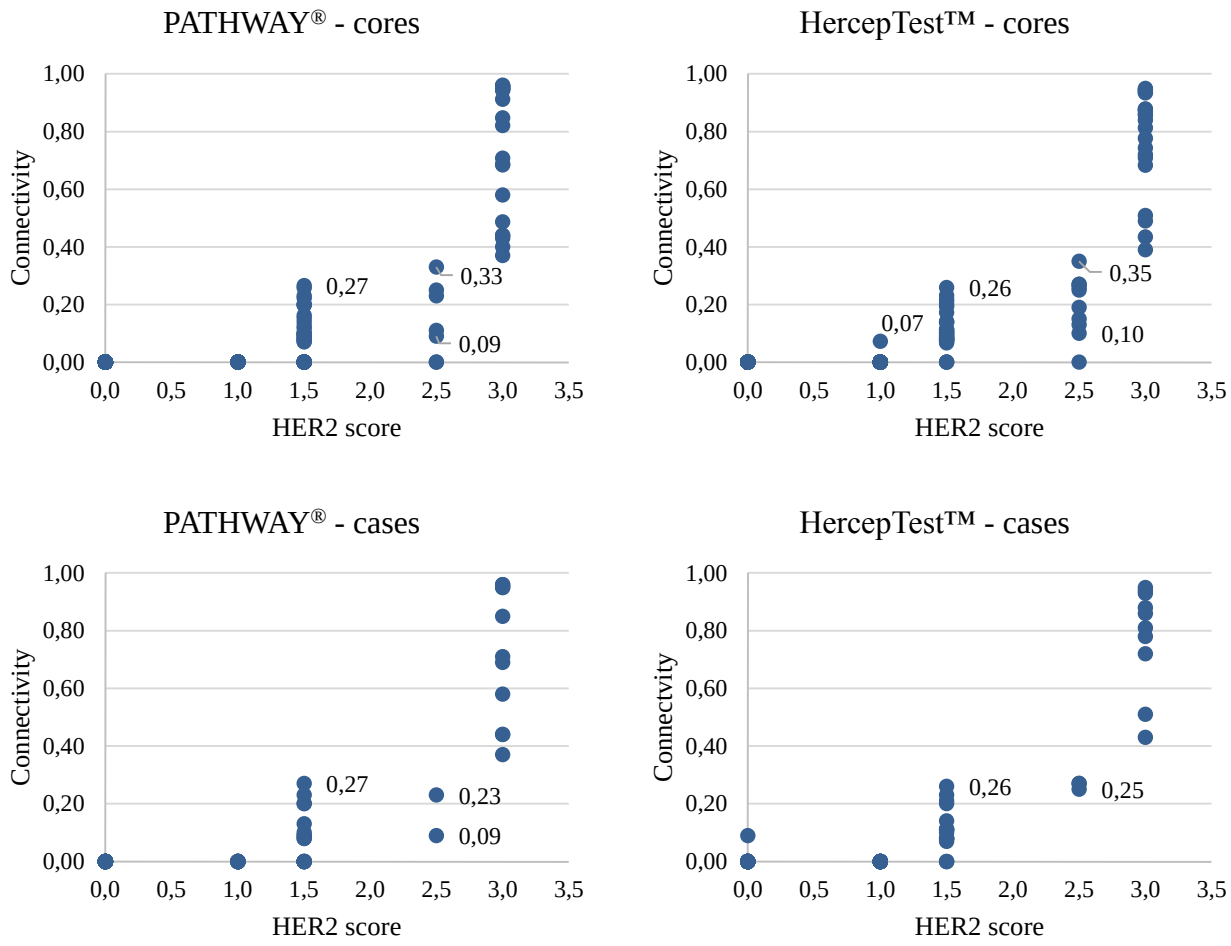

Figure 4: Correlation between connectivity and HER2 score for cores and cases

The lowest connectivity value for a core with a 2.5 HER2 score (HER2 IHC 2+ / FISH pos) was 0.09 for PATHWAY® and 0.10 for HercepTest™. This value changed for HercepTest™ to 0.25, when the highest connectivity value from the three cores per case was selected to represent the case. For PATHWAY®, 0.09 was still the highest value per case for a 2.5 HER2 score.

Two cores for PATHWAY® and one core for HercepTest™ had a HER2 score of 2.5, but a connectivity of 0.00. The two other cores from the associated cases had higher connectivity values, classifying the cases as either equivocal or positive. From figure 4, adjusted levels for the cut-offs were obtained:

Negative: 0.00-0.08

Equivocal: 0.09-0.30

Positive: >0.30

Connectivity values were converted to DIA scores, applying the altered cut-off values.

The new HER2 CONNECT™ results were compared to FISH (table 8).

Table 8: Comparison of HER2 CONNECT™ results to FISH applying altered cut-off values

PATHWAY® Cases	FISH		Total	HercepTest™ Cases	FISH		Total
	neg	pos			neg	pos	
Negative	74	0	74	**Negative**	78	0	78
Equivocal	8	2	10	**Equivocal**	11	3	14
Positive	0	11	11	**Positive**	0	10	10
Total	82	13	95	**Total**	89	13	102
PATHWAY® Cores	**FISH**		**Total**	**HercepTest™ Cores**	**FISH**		**Total**
	neg	**pos**			**neg**	**pos**	
Negative	243	2	245	**Negative**	254	1	255
Equivocal	17	4	21	**Equivocal**	20	8	28
Positive	0	23	23	**Positive**	0	22	22
Total	260	29	289	**Total**	274	35	305

When the adjusted cut-off values were applied, 74 cases were classified as negative, 10 cases as equivocal and 11 cases as positive for PATHWAY®. For HercepTest™, 78 cases were classified as negative, 14 as equivocal and 10 as positive. All cases classified as positive by HER2 CONNECT™ were amplified by FISH. Two of 10 equivocal cases for PATHWAY® and three of 14 cases for HercepTest™ had HER2 gene amplification. All cases classified as negative by HER2 CONNECT™ were non-amplified.

False negative results occurred when the individual cores instead of cases were compared to FISH. For PATHWAY® and HercepTest™ two and one cores, respectively, were scored as negative by DIA, but amplified by FISH. The cores were all scored as 2+ by manual IHC interpretation.

The sensitivity, specificity and accuracy was 100% for both PATHWAY® and HercepTest™, when the cases were used for calculation.

DIA compared to manual IHC interpretation

Distribution of scores from manual IHC interpretation was compared to results from HER2 CONNECT™ analysis (table 9)

Table 9: Distribution of scores according to manual IHC interpretation and HER2 CONNECT™ analysis of cases

Score	PATHWAY® IHC	PATHWAY® HER2 CONNECT™
0/1+	66	76
2+	20	10
3+	11	11
Total	97	97
	HercepTest™ IHC	**HercepTest™ HER2 CONNECT™**
0/1+	72	80
2+	22	14
3+	10	10
Total	104	104

The number of equivocal cases was reduced by 10 cases for PATHWAY® and 8 cases for HercepTest™. The reduction equaled 50.0% and 36.4%, respectively, which were statistically significant ($p < 0.05$) for both assays. There was total agreement regarding 3+ cases.

Comparison of HER2 CONNECT™ results for PATHWAY® and HercepTest™

The agreement between PATHWAY® and HercepTest™ was determined by comparing all cores. The agreement was analyzed by calculation of Cohen's Kappa with use of DIA scores derived from connectivity. A kappa value of 0.79 was found, indicating substantial agreement[25].

Discussion

Assessment of HER2 status in gastric and GEJ adenocarcinomas is essential to identify patients, who are candidates for treatment with trastuzumab. An accurate method is mandatory and a cost-effective approach is highly appreciable. This study is to our best knowledge the first to evaluate the accuracy of HER2 CONNECT™ in relation to gastric and GEJ adenocarcinomas.

In our study, the prevalence of HER2 gene amplification was 14.4%, which correlates with reported prevalence rates in the litterature[6]. Two of the HER2 gene amplified cases were scored as negative by IHC. This phenomenon has previously been reported[10]. The cases were excluded from HER2 CONNECT™ analysis, as they would not

permit treatment with trastuzumab according to EMA[8].

HER2 CONNECT™ was originally developed as a diagnostic tool for breast cancer samples. Cut-off values, which convert connectivity to an equivalent IHC score, have been determined for breast carcinoma. Applying these cut-off values to our study material resulted in three HER2 amplified cases being classified as negative by DIA. All three cases were scored as equivocal by manual IHC interpretation, except one case which was scored as positive by PATHWAY®.

Adjustment of the cut-off values was a possible approach to enhance precision of HER2 CONNECT™, when analyzing gastric and GEJ adenocarcinomas. The main focus for altering the cut-offs were to minimize the risk of false negative results and to reduce the proportion of equivocal cases requiring additional ISH for final classification. The altered cut offs were determined and defined a connectivity of < 0.09 as negative, 0.09-0.30 as equivocal and > 0.30 as positive. Applying the altered cut-offs resulted in a 100% sensitivity and specificity for both IHC assays. A statistically significant reduction of equivocal cases of 50.0% for PATHWAY® and 36.4% for HercepTest™ was observed for HER2 CONNECT™. The cut-off values determined in this study need further validation and confirmation with inclusion of different samples i.e. whole-sections, biopsy specimens and samples from different institutions.

Comparison of connectivity with HER2 score, revealed that the lowest connectivity value for a case with a 2.5 HER2 score (IHC 2+ and amplified by FISH) was 0.09 for PATHWAY® and 0.25 for HercepTest™. Substantial agreement ($\kappa = 0.79$) was found between the IHC assays for results from HER2 CONNECT™ analysis. However, the difference in lowest connectivity for a 2.5 HER2 score indicate that separate cut-offs for different IHC assays could enhance precision of HER2 CONNECT™. More precisely, the cut-off interval (0.09-0.30) classifying cases as

equivocal could be downsized for HercepTest™ to further enhance accuracy of the software. However, individually determined cut-off values for different IHC assays would require several validation runs for each assay to be used by the software.

False negative results occurred when DIA scores for individual cores, instead of cases, were compared to FISH results. For PATHWAY® and HercepTest™, two and one cores, respectively, were classified as false negative. The cores were scored as 2+ by manual IHC interpretation. The other cores from the respective cases classified the case as equivocal or positive. This might be a potential challenge for HER2 CONNECT™ with the present software configuration for cases with equivocal HER2 protein expression.

Connectivity and manual IHC score was discrepant in some cases, especially when the HER2 expression was heterogeneous within a single core. The discrepant cases were reanalyzed and regions of interest (ROI) were manually selected. It was observed that the highest connectivity value was achieved when ROI only included the area with the strongest HER2 expression.

The HER2 CONNECT™ software recognized artifacts (pigments etc.) as membranous staining, which lead to falsely high connectivity values. These cases were easily identified because of discrepancy between connectivity and manual IHC score and the DIA analysis evidently was based on non-relevant structures.

Cytoplasmic and nuclear staining was observed for PATHWAY®. The intensity of the non-specific staining was frequently strong, interfering with interpretation. Seven cases were scored as equivocal for this reason and were excluded from HER2 CONNECT™ analysis. The software registered cytoplasmic and nuclear staining as membranous which gave falsely high connectivity values. The reduction of equivocal 2+ IHC cases was higher for PATHWAY® compared to HercepTest™, 50.0% vs. 36.4% when HER2 CONNECT™ analysis were performed.

However, the seven cases scored as equivocal because of non-specific staining still needed retesting by FISH. Therefore, a total of 17 cases were scored as equivocal by HER2 CONNECT™ analysis with PATHWAY®, compared to 14 for HercepTest™. HercepTest™ in combination with HER2 CONNECT™ may be a superior method compared to PATHWAY® as all cases could be evaluated with reduction of equivocal cases and 100% sensitivity.

The HER2 CONNECT™ software algorithm was directly applied to gastric and GEJ cancer, only altering the cut-off values. Adjustment of the algorithm itself may enhance the effectiveness of the software. It might further reduce the number of equivocal cases and increase precision when individual cores are compared to FISH. It is unknown whether alteration of the software algorithm could enable the software to distinguish the non-specific staining from membranous, which was a challenge for IHC performed with PATHWAY®. The cut-off values for gastric and GEJ cancer, set in this study, was lower than cut-offs determined for breast cancer. One of the reasons for this difference could be the incomplete staining of membranes, which is characteristic for gastric and GEJ cancers. The software eliminates small membrane fragments from calculation of connectivity according to a specified cut-off size. This cut-off might be changed to optimize the compatibility of HER2 CONNECT™ to gastric and GEJ cancers.

Approximately 15 % of the cases included in this study were classified as heterogeneous (different IHC scores for cores from the same case), indicating the importance of including multiple tumor cores for a TMA set-up. To our knowledge, only two studies[22,26] have investigated the concordance rates between TMAs and whole sections in relation to gastric and GEJ adenocarcinomas. Machado et al.[26] used two tissue cores of 0.6 mm to represent the resected tumor and found a poor concordance to whole sections. Gasljevic et al.[22] also used two tissue cores for each

resection specimen, but the size of the cores was 2.0 mm. This study found an overall concordance rate of 85%. More studies are needed to further investigate how many tissue cores are needed to represent a tumor when applying a TMA approach for gastric and GEJ adenocarcinomas. Heterogeneous cases typically reveal small foci of tumor cells with HER2 protein overexpression. These foci can be difficult to identify during FISH analysis. A combined gene protein assay has been developed by Ventana to ease interpretation. The assay allows simultaneous interpretation of HER2 protein expression and HER2 gene amplification in the same slide. Few studies have evaluated the assay for gastric and GEJ adenocarcinomas, exhibiting promising results[27-29].

In conclusion, this study has shown that HER2 CONNECT™ seems to be an accurate tool in assessment of HER2 expression in gastric and GEJ cancer. Adjustment of cut-off values determined for breast cancer ensured a 100% analytical sensitivity, specificity and accuracy when HER2 CONNECT™ was applied. A statistically significant reduction of equivocal cases for PATHWAY® (50.0%) and HercepTest™ (36.4%) was achieved for HER2 CONNECT™. Adjustment of the software can potentially enhance the precision of HER2 CONNECT™. Further studies with inclusion of additional samples types are required to validate results from this study.

References

1. Ferlay J, Soerjomataram I, Ervik M, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase no. 11. Lyon, France: International Agency for Research on Cancer; 2013. http://globocan.iarc.fr. Updated 2013. Accessed 03/09, 2015.

2. Howlader N, Noone A, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2011, National Cancer Institute. Bethesda, MD, based on November 2013 SEER data submission. http://seer.cancer.gov/csr/1975_2011/. Updated 2014. Accessed 03/09, 2015.

3. Tan YK, Fielding JWL. Early diagnosis of early gastric cancer. *Eur J Gastroenterol Hepatol*. 2006;18(8):821-829.

4. Howlader N, Noone A, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD, based on November 2014 SEER data submission. http://seer.cancer.gov/csr/1975_2012/. Updated April 2015. Accessed 08/12, 2015.

5. Carcas LP. Gastric cancer review. *J Carcinog*. 2014;13:14-3163.146506. eCollection 2014.

6. Jørgensen JT. Role of human epidermal growth factor receptor 2 in gastric cancer: Biological and pharmacological aspects. *World J Gastroenterol*. 2014;20(16):4526-4535.

7. Hudis CA. Trastuzumab — mechanism of action and use in clinical practice. *N Engl J Med*. 2007;357(1):39-51.

8. Europeans Medicines Agency. Assesment report for herceptin. procedure no. EMEA/H/C/278/II/0047 http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000278/WC500074921.pdf Updated 2010. Accessed 08/12, 2015.

9. U.S. Food and Drug Administration. Trastuzumab. http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm230418.htm. Updated 2010. Accessed 08/12, 2015.

10. Bang Y, Van Cutsem E, Feyereislova A, et al. 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. *The Lancet*. 2010;376(9742):687-97.

11. Hofmann M, Stoss O, Shi D, et al. Assessment of a HER2 scoring system for gastric cancer: Results from a validation study. *Histopathology*. 2008;52(7):797-805.

12. Rüschoff J, Dietel M, Baretton G, et al. HER2 diagnostics in gastric cancerâ€"guideline validation and development of standardized immunohistochemical testing. *Virchows Archiv*. 2010;457(3):299-307.

13. NordiQC. Assesment Run G2 2011 (gastric cancer pilot module). http://www.nordiqc.org/Run-33-B12-G2/Assessment/assessment-G2-HER2.htm. Updated 2012. Accessed 08/24, 2015.

14. UK NEQAS ICC & ISH. Run 109: The gastric HER2 ICC module. http://www.ukneqasiccish.org/wp/wp-

content/uploads/2015/07/run_109_journal.pdf Updated 2015. Accessed 08/24, 2015.

15. 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(1):18-43.

16. Brugmann A, Eld M, Lelkaitis G, et al. Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. *Breast Cancer Res Treat*. 2012;132(1):41-49.

17. Radu OM, Foxwell T, Cieply K, et al. HER2 amplification in gastroesophageal adenocarcinoma: Correlation of two antibodies using gastric cancer scoring criteria, H score, and digital image analysis with fluorescence in situ hybridization. *Am J Clin Pathol*. 2012;137(4):583-594.

18. Jeung J, Patel R, Vila L, Wakefield D, Liu C. Quantitation of HER2/neu expression in primary gastroesophageal adenocarcinomas using conventional light microscopy and quantitative image analysis. *Arch Pathol Lab Med*. 2012;136(6):610-617.

19. Fusco N, Rocco EG, Del Conte C, et al. HER2 in gastric cancer: A digital image analysis in pre-neoplastic, primary and metastatic lesions. *Mod Pathol*. 2013;26(6):816-824.

20. Cho EY, Srivastava A, Park K, et al. Comparison of four immunohistochemical tests and FISH for measuring HER2 expression in gastric carcinomas. *Pathology*. 2012;44(3):216-220.

21. Sheffield BS, Garratt J, Kalloger SE, et al. HER2/neu testing in gastric cancer by immunohistochemistry: Assessment of interlaboratory variation. *Arch Pathol Lab Med*. 2014;138(11):1495-1502.

22. Gasljevic G, Lamovec J, Contreras JA, Zadnik V, Blas M, Gasparov S. HER2 in gastric cancer: An immunohistochemical study on tissue microarrays and the corresponding whole-tissue sections with a supplemental fish study. *Pathol Oncol Res*. 2013;19(4):855-865.

23. VISIOPHARM. HER2-CONNECT. http://visiopharm.com/appcenter/?id=10007. Accessed 11/24, 2015.

24. Holten-Rossing H, Moller Talman ML, Kristensson M, Vainer B. Optimizing HER2 assessment in breast cancer: Application of automated image analysis. *Breast Cancer Res Treat*. 2015;152(2):367-375.

25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.

26. Abrahao-Machado LF, Jacome AA, Wohnrath DR, et al. 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.

27. Werner D, Battmann A, Steinmetz K, et al. The validation of a novel method combining both HER2 immunohistochemistry and HER2 dual-colour silver in situ hybridization on one slide for gastric carcinoma testing. *J Transl Med*. 2014;12:160-5876-12-160.

28. Hirschmann A, Lamb TA, Marchal G, Padilla M, Diebold J. Simultaneous analysis of HER2 gene and protein on a single slide facilitates HER2 testing of breast and gastric carcinomas. *Am J Clin Pathol*. 2012;138(6):837-844.

29. Nishida Y, Kuwata T, Nitta H, et al. A novel gene-protein assay for evaluating HER2 status in gastric cancer: Simultaneous analyses of HER2 protein overexpression and gene amplification reveal intratumoral heterogeneity. *Gastric Cancer*. 2015;18(3):458-466.