Assessment of a HER2 scoring system for colorectal cancer: results from a validation study

Emanuele Valtorta[1,19], Cosimo Martino[2,19], Andrea Sartore-Bianchi[1], Frédérique Penaullt-Llorca[3], Giuseppe Viale[4], Mauro Risio[2], Massimo Rugge[5], Walter Grigioni[6], Katia Bencardino[1], Sara Lonardi[7], Vittorina Zagonel[7], Francesco Leone[2], Johannes Noe[8], Fortunato Ciardiello[9], Carmine Pinto[6], Roberto Labianca[10], Stefania Mosconi[10], Claudio Graiff[11], Giuseppe Aprile[12], Barbara Frau[13], Carlo Garufi[14], Fotios Loupakis[15], Patrizia Racca[16], Giuseppe Tonini[17], Calogero Lauricella[1], Silvio Veronese[1], Mauro Truini[1], Salvatore Siena[1,18,20], Silvia Marsoni[2,20] and Marcello Gambacorta[1,20]

[1]Niguarda Cancer Center, Ospedale Niguarda Ca' Granda, Milano, Italy; [2]Istituto di Candiolo, FPO-IRCCS, Candiolo, Italy; [3]Centre Jean Perrin, University of Auvergne, Clermont-Ferrand, France; [4]Istituto Europeo Oncologia, Milano, Italy; [5]Anatomia Patologica, Azienda Ospedale-Università di Padova, Padova, Italy; [6]Policlinico Sant'Orsola-Malpighi, Bologna, Italy; [7]Oncologia Medica 1, Istituto Oncologico Veneto – IRCCS, Padova, Italy; [8]F. Hoffmann-La Roche, Basel, Switzerland; [9]Seconda Università di Napoli, Napoli, Italy; [10]Ospedale Papa Giovanni XXIII, Bergamo, Italy; [11]Ospedale di Bolzano Via Boehler, Bolzano, Italy; [12]Ospedale Santa Maria della Misericordia, Piazzale Santa Maria della Misericordia, Udine, Italy; [13]Policlinico Universitario di Cagliari, Monserrato, Italy; [14]Istituto Nazionale Tumori Regina Elena, Roma, Italy; [15]Azienda Ospedaliero-Universitaria Pisana Via Roma, Pisa, Italy; [16]Azienda Ospedaliero-Universitaria Città della Salute e della Scienza, Torino, Italy; [17]Policlinico Universitario Campus Biomedico, Roma, Italy and [18]Università degli Studi di Milano, Milano, Italy

We sought to develop criteria for *ERBB2*-positivity (*HER2*) in colorectal cancer to ensure accurate identification of *ERBB2*-amplified metastatic colorectal cancer patients suitable for enrolment in a phase II trial of ERBB2-targeted therapy (HERACLES trial). A two-step approach was used. In step 1, a consensus panel of pathologists adapted existing protocols for use in colorectal cancer to test *ERBB2* expression and amplification. Collegial revision of an archival test cohort of colorectal cancer samples led to specific recommendations for adapting current breast and gastric cancer criteria for scoring *ERBB2* in colorectal cancer. In step 2, from September 2012 to January 2015, colorectal-specific *ERBB2* testing protocols and *ERBB2* scoring criteria were used to centrally screen for *ERBB2*-positive *KRAS* wild-type colorectal cancer patients to be enrolled in the HERACLES trial (clinical validation cohort). In both archival test ($N = 256$) and clinical validation ($N = 830$) cohorts, a clinically sizeable 5% fraction of *KRAS* wild-type colorectal cancer patients was found to be *ERBB2*-positive according to the colorectal cancer-specific *ERBB2* scoring criteria. *ERBB2*-positive tumors showed ERBB2 immunostaining consisting of intense membranous ERBB2 protein expression, corresponding to homogenous *ERBB2* amplification, in >50% of cells. None of the immunohistochemistry 0 or 1+ cases was amplified. Concordance between SISH and FISH was 100%. In conclusion, we propose specific criteria for defining *ERBB2*-positivity in colorectal cancer (HERACLES Diagnostic Criteria). In a phase II trial of trastuzumab and lapatinib in a cetuximab-resistant population, HERACLES Diagnostic Criteria shaped the selection of patients and defined *ERBB2* as a predictive marker for response to ERBB2-targeted therapy in metastatic colorectal cancer.
Modern Pathology (2015) **28**, 1481–1491; doi:10.1038/modpathol.2015.98; published online 9 October 2015

The human epidermal growth factor receptor 2 gene *ERBB2* (*HER2*) is amplified and its protein is over-expressed in many cancer types.[1,2] Overexpression of ERBB2 is an established therapeutic target in breast and gastric cancers and is successfully exploited in the clinic using a variety of anti-ERBB2 agents, leading to remarkable outcome improvements.[3,4] Although the comprehensive molecular characterization of human colorectal cancer has identified *ERBB2* amplification as a potential therapeutic target[5] and ERBB2 over-expression has been controversially linked to prognosis,[6–9] the clinical significance of *ERBB2* alterations remains elusive. Recently, we and others have found that activation of ERBB2 signaling causes resistance to anti-EGFR therapy in a fraction of metastatic colorectal patients, wild type for *RAS* codons 12–13.[10–12] Of more relevance for the clinic, we have demonstrated that the combination of the anti-ERBB2 monoclonal antibody trastuzumab and the dual EGFR/HER2 small-molecule inhibitor lapatinib, but not either drug alone, is effective in inducing durable tumor shrinkage in *ERBB2*-amplified meta-static CRC patient-derived xenografts.[10]

Reasoning that *ERBB2* corsivo could also represent a valuable therapeutic target in *KRAS* wild-type metastatic colorectal cancer patients resistant to anti-EGFR treatment, we designed the HERACLES trial, a phase II trial testing the combination of trastuzumab and lapatinib in *ERBB2*-positive metastatic colorectal cancer patients refractory to standard treatment, including cetuximab or panitumumab.[11] Prior to starting the HERACLES trial, we elected to develop colorectal cancer-specific criteria for the definition of *ERBB2* positivity. Immunohistochemistry and fluorescent or silver *in situ* hybridization (FISH or SISH) are current standard methodologies to detect respectively ERBB2 protein expression and gene amplification on formalin-fixed paraffin-embedded tumor samples. These are routinely used to establish *ERBB2* status in breast and gastric cancer but have not been customized for colorectal cancer for which the reported rate for *ERBB2* positivity ranges enormously from < 1% to > 50%.[6–9,12–20] The aim of the present study was to develop a validated ERBB2 scoring system for colorectal cancer with the goal of identifying *ERBB2*-positive patients suitable for enrollment in the HERACLES trial.

Materials and methods

Study Design

This was a two-steps study. Step 1 was conducted on formalin-fixed paraffin-embedded archival samples (archival test cohort). A consensus panel of three pathologists (i) defined, by similarity with breast and gastric cancers, the technical protocols of assessment for two immunohistochemistry staining protocols and two *in situ* hybridization methods; (ii) appointed one member of the consensus panel (MG)

to review all samples; (iii) established the criteria by which samples for collegial review were to be selected by the appointed pathologist for consensus revision; (iv) collegially read and discussed the characteristics of the selected samples during a day-long consensus session, and (v) as a result of the consensus review formulated a diagnostic algorithm for *ERBB2* positivity in colorectal cancer, referred to as HERACLES Diagnostic Criteria. Step 2 was conducted on samples from *KRAS* 12/13 wild-type metastatic colorectal cancer patients prospectively screened for the HERA-CLES trial (clinical validation cohort). A centralized pathology laboratory (Niguarda Cancer Center, Milan, Italy) processed all samples, including those already tested for ERBB2 at HERACLES participating centers. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice. Informed consent, allowing the use of the patient's surgical/bioptical specimen for diagnostic purposes, was available for each archival sample of the archival test cohort, while for the screening (clinical validation cohort), patients signed a protocol-specific informed consent approved by independent Ethical Committees.

ERBB2 Status

ERBB2 expression analysis by immunohistochemistry was performed manually using HercepTest antibody (Dako A/S Glostrup, Denmark) and automatically on the automated Bench Mark Ultrasystem using the VENTANA 4B5 antibody, following the manufacturers' instructions in both cases.

ERBB2 amplification analysis by FISH was performed with a PathVysion HER-2 DNA Probe Kit (Abbott Laboratories, Des Plaines, IL, USA) and SISH with a VENTANA 4B5 Inform HER2 dual-color on the BenchMark Ultra system (Inform HER2 DNA dual-color assay—Roche Tissue Diagnostics, VENTANA Medical Systems, SA). The scoring and evaluation for *in situ* hybridization was performed by counting *ERBB2* and *CEN17* signals from 100 nuclei per case. Non-tumor tissue (normal colon mucosa) was used as an internal negative control. Samples with a *ERBB2*/*CEN17* ratio ≥ 2.0 were considered amplified. Images were captured with the Axiovision software using an Axio Zeiss Imager 2 microscope for IHC and SISH and the ISIS Metasystems software using an Axio Zeiss Imager Z1 microscope for FISH.

Analysis

A sample was considered evaluable for review when all test results were present. At least one evaluable sample was required to consider a patient evaluable. In the case of multiple samples, the highest immuno-histochemistry value defined the patient score. In step 1, scoring procedures for immunohistochemistry and *in situ* hybridization took into account the standard scoring systems for breast and gastric cancers.[21–23] In the first analysis, three staining parameters, ie, pattern

HER2 assessment in colorectal cancer
E Valtorta *et al*
1483
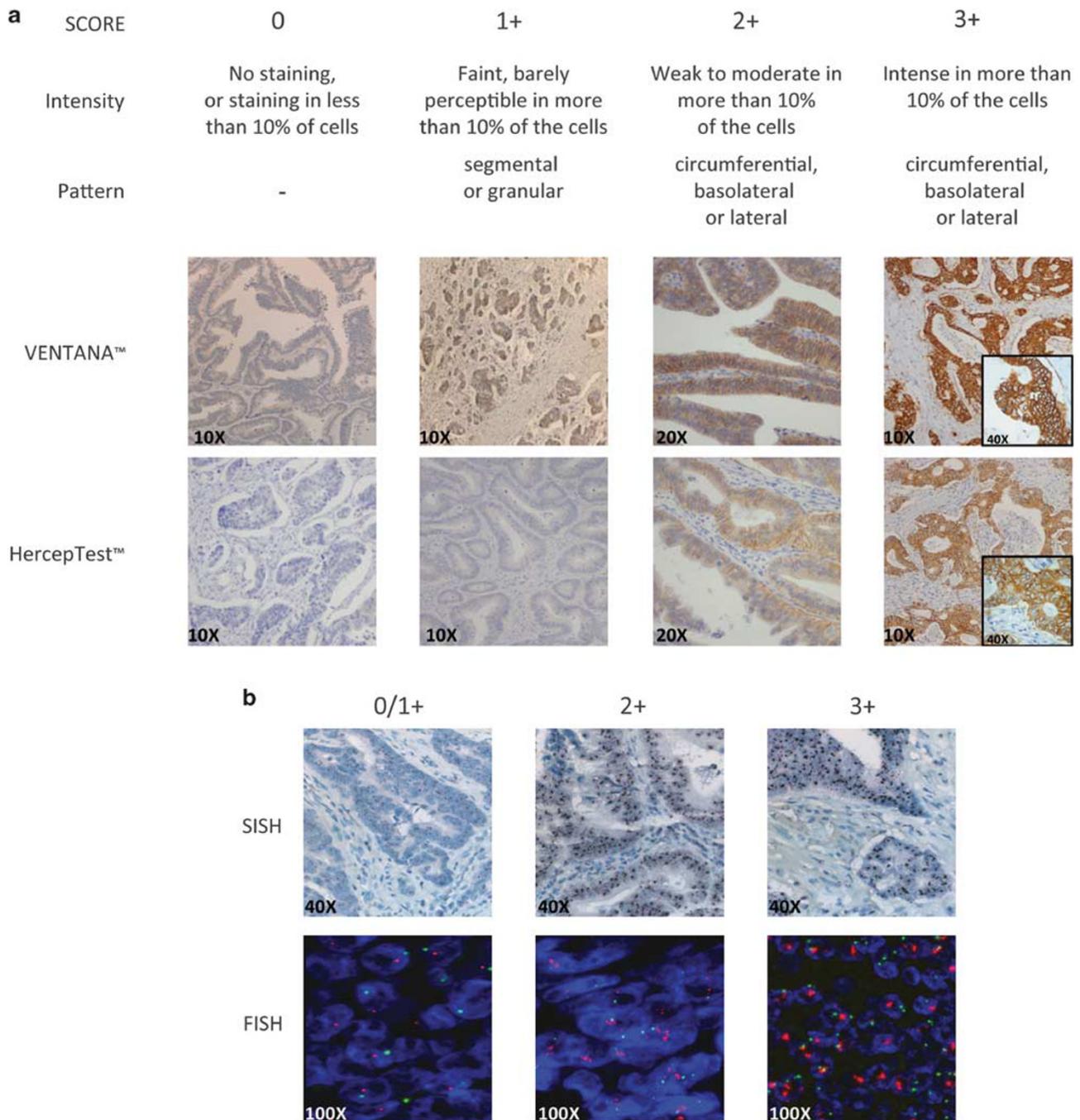
Figure 1 Initial criteria for ERBB2 determination of immunohistochemistry scores (0/1+/2+/3+) (**a**); photomicrography of typical ERBB2 protein expression (VENTANA 4B5 and HercepTest) (**b**); and *ERBB2* gene amplification (fluorescent (FISH) or silver *in situ* hybridization (SISH)) (**c**) in colorectal cancer archival samples.

of membranous reactivity, intensity of reactivity, and percentage of immunoreactive cells, were combined as shown in Figure 1a. Only samples with a ratio *ERBB2*: *CEN17* ≥ 2 or staining equivocal (2+) or positive (3+) in at least 10% of cells were selected for review, together with a small set of samples scoring immuno-histochemistry 0/1+ as negative controls. In step 2, only the VENTANA 4B5 and FISH kits were used for

ERBB2 determination. All samples were centrally scored according to HERACLES Diagnostic Criteria.

We used percentage and mean for qualitative variables and s.d. for quantitative variables. Receiver operating characteristic (ROC) curves were used to determine test performances for each IHC method; accuracy, sensitivity, specificity, and positive and negative predictive values, were calculated in the

two series using *in situ* hybridization as the gold standard and a cutoff for immunohistochemistry positivity as defined by the HERACLES Diagnostic Criteria.

Results

Archival Test Cohort (Step 1)

The characteristics of the 359 patients with colorectal adenocarcinoma who provided 482 tumor samples for the archival test cohort and the Study Consort are reported in Supplementary Figures S1A and B. Of the 482 samples, 66 (14%, derived from 49 patients) were from *KRAS* exon 2 mutated cancers. Overall, 134 samples (22 *KRAS* mutated) could not be assessed with all planned procedures because of insufficient tumor material ($N = 36$, 27%), initial technical problems with the first SISH method used (INFORM HER2-VENTANA, $N = 94$, 70%) and sub-optimal tissue preparation ($N = 4$, 3%), leaving 348 evaluable samples (44 *KRAS* mutated) (Supplementary Figure S1B). Fifty-eight percent of the evaluable *KRAS* wild-type samples were from primary cancers, 13% from metastatic sites, and 29% from both primary and metastatic sites (Supplementary Figures S1C and D).

Immunohistochemistry and SISH results are reported in Table 1. Of the 44 *KRAS* mutated samples (Table 1A), none was *ERBB2* amplified or expression positive (3+). Equivocal (2+) staining was observed in six samples with VENTANA 4B5 and one sample with HercepTest, leaving 97% HercepTest and 86% VENTANA 4B5 negative samples. In KRAS wild-type samples (Table 1B), immunohistochemistry positivity (3+) with strong circumferential membrane immunostaining was observed in 14 cases (4.6%) with HercepTest and 16 cases (5.3%) with VENTANA 4B5. Equivocal staining (2+) with weak-to-moderate circumferential basolateral or lateral immunoreactivity was seen in >10% of cells and was more frequent with

VENTANA 4B5 ($N = 23$, 7.6%) than with HercepTest ($N = 3$, 1%). VENTANA 4B5 did not stain (0) or faintly stained (1+) 87% of samples (0, $N = 170$; 1+, $N = 95$; total $N = 265$); none of these samples was amplified by SISH. HercepTest stained negative (0) in 74% of samples ($N = 226$), including one SISH-positive sample (staining 3+ with VENTANA 4B5) and faintly stained (1+) the remaining 61 samples (20%), none of which was amplified. The comparative analysis between immunohistochemistry and SISH results, used as the 'gold standard', is also reported in Table 1. Overall, *ERBB2* amplification in *KRAS* wild-type cases was observed in 17 samples (5.6% of samples, corresponding to 5.1% of patients), and of these 16 (94.1%) scored positive (3+) and 1 equivocal (2+) with VENTANA 4B5. With HercepTest, 14 (82.4%) were positive (3+), 2 equivocal (2+) and 1 negative (0).

Concordance between samples from primary and metastatic tumor sites in the same patients was possible for 95 paired samples from 47 patients (1 patient had two metastatic site samples, Supplementary Table S1). There were four *ERBB2* amplified cases (8.5%): VENTANA 4B5 stained all four paired samples positive (3+), while HercepTest only three of the pairs. The remaining amplified case scored equivocal (2+) in the primary and predominantly negative (0) in the paired liver metastasis. This latter case, however, was also the only one showing clear intra-sample heterogeneity out of the 17 *ERBB2*-amplified cases, with areas of strong 3+ reactivity to VENTANA 4B5 coexisting with areas of no (0) or faint reactivity (1+) corresponding to *ERBB2*-amplified and non-amplified tumor areas, respectively, in both the primary and the metastatic tumors (Figure 2).

Consensus Panel of Pathologists Review

The Consensus Panel of Pathologists reviewed only cases that were informative for all three analyses.

Table 1 Results of the immunoistochemistry test with HercepTest or Ventana 4B5 and silver *in situ* hybridization (SISH)

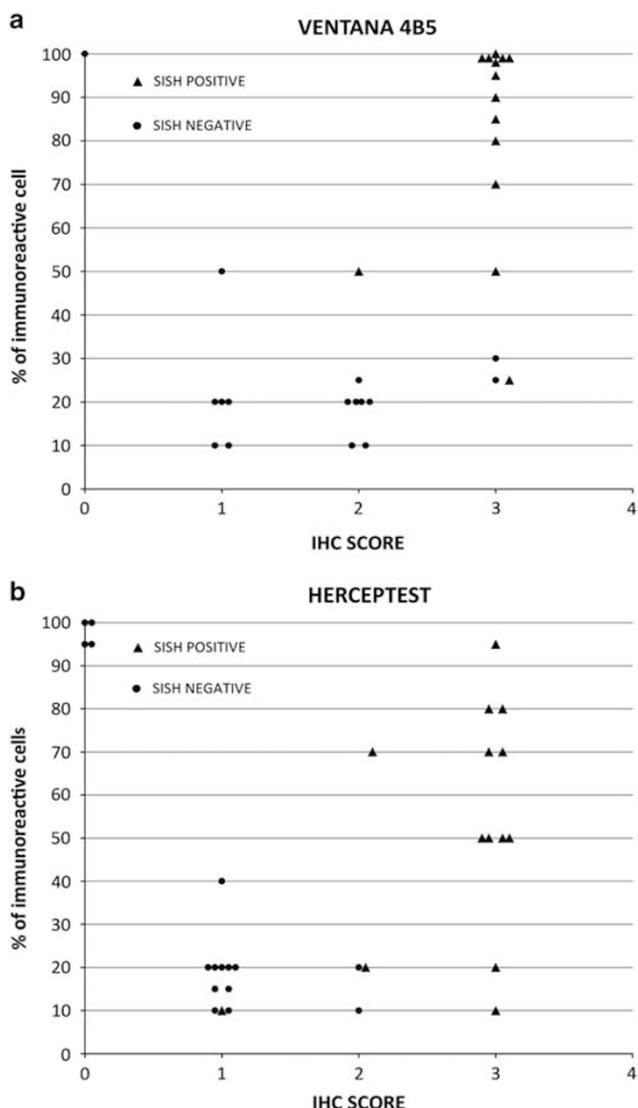|	HercepTest 3+	HercepTest 2+	HercepTest 1+	HercepTest 0	Total	SISH amplified
(A) Archival cohort: *KRAS* exon 2 (codon 12/13) mutated samples						
VENTANA 3+	0	0	0	0	0	—
VENTANA 2+	0	1	4	1	6	—
VENTANA 1+	0	0	9	14	23	—
VENTANA 0	0	0	2	13	15	—
Total	0	1	15	28	44	—
SISH amplified	—	—	—	—	—	0/44
(B) Archival cohort: *KRAS* exon 2 (codon 12/13) WT samples						
VENTANA 3+	<u>14</u>	<u>1</u>	0	<u>1</u>	<u>16</u>	94.1%
VENTANA 2+	0	<u>1</u>	10	12	23 (<u>1</u>)	5.9%
VENTANA 1+	0	1	43	51	95	—
VENTANA 0	0	0	8	162	170	—
Total	<u>14</u>	3 (<u>2</u>)	61	226 (<u>1</u>)	304 (<u>17</u>)	100.0%
Percentage of amplified	82.4%	11.8%	—	5.9%	100.0%	<u>17</u>/304 (5.6%)

Amplified samples are underlined.
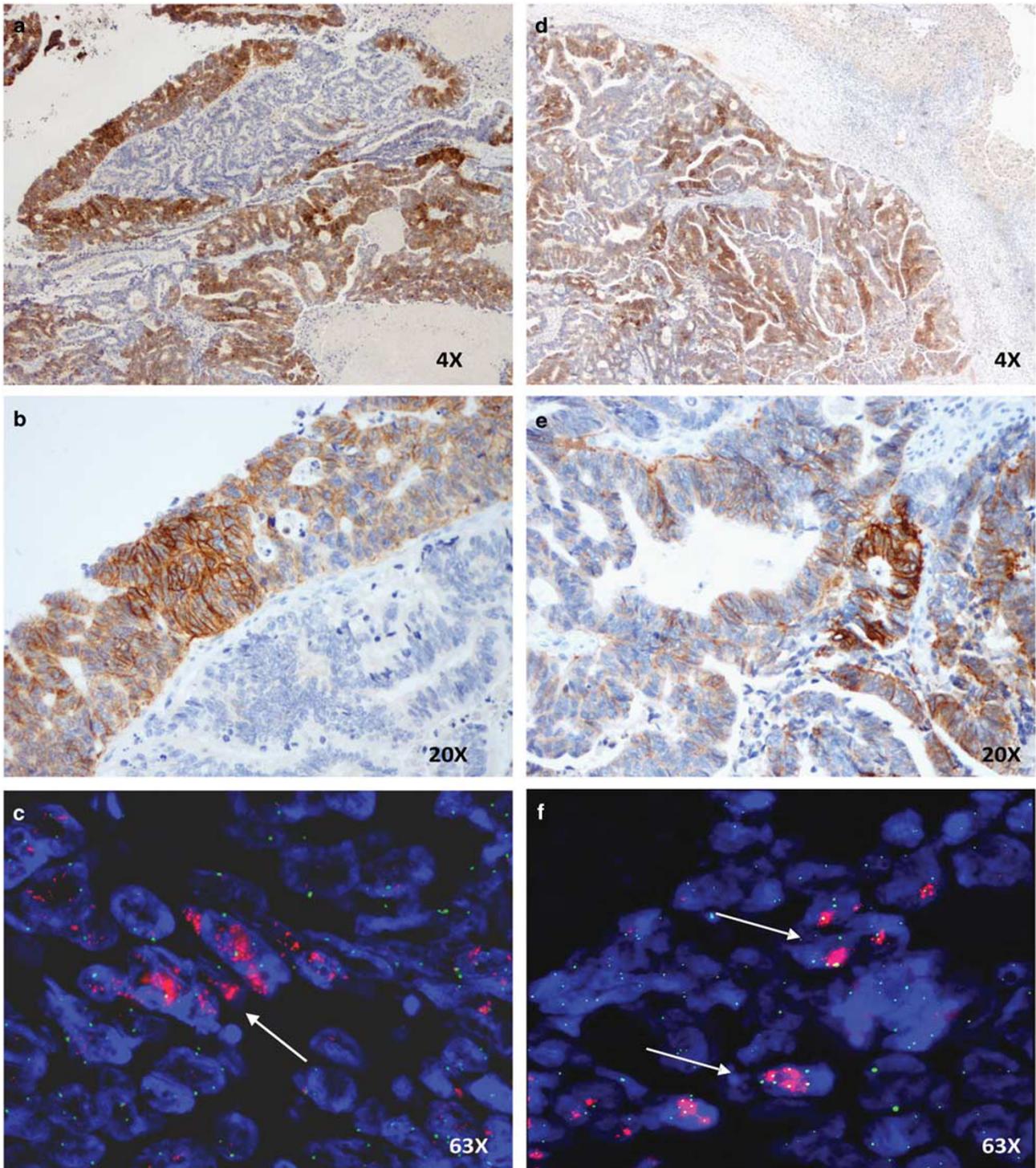
Figure 2 Atypical pattern of heterogeneous ERBB2 expression and amplification in colorectal cancer (**a–c**) and liver metastasis (**d–f**) from a single patient (#01114) clearly shows clusters of tumor cells with different immunoreactivity ranging from 0 to 3+, corresponding to areas of high (arrows) and no amplification. (**a**, **b**, **d**, **e**) ERBB2 protein expression by VENTANA 4B5 at x4 (**a**, **d**) and x20 (**b**, **e**). (**c**, **f**) *ERBB2* gene copy number by fluorescent *in situ* hybridization at x63 magnification.

Although a formal technique, such as Delphi, was not used to reach consensus, the three pathologists reviewed 30 positive (3+) or equivocal (2+) and a random selection of 6 0/1+ cases, derived in total from 25 patients; these were read at multi-head microscopes and discussed until a consensus was reached. According to the Consensus Panel of Pathologists, all *ERBB2*-amplified cases examined (*N* = 14) showed a typical immunoreactivity pattern consisting of circumferential or basolateral or lateral immunohistochemistry staining of the cancer cells, resulting in highly homogenous and intensely

Figure 3 Archival test cohort: consensus panel results: scatter plot of the immunohistochemistry score and percentage of immunoreactive cells obtained with VENTANA 4B5 (**a**) or HercepTest (**b**) versus amplification assessed with silver *in situ* hybridization (SISH; black triangles).

stained tumor areas, with normal mucosa or liver tissue background of moderate intensity. Cytoplasmic staining of the cancer cells was not considered. The Consensus Panel of Pathologists agreed that the percentage of immunoreactive tumor cells within each sample significantly differed between VENTANA 4B5 and HercepTest (Figures 3a and b). With VENTANA 4B5, 93% (N = 13) of the 14 amplified samples scored immunohistochemistry positive in ≥50% of cells, mostly clustering at or above the 90% cellularity mark, while non-amplified positive (3+) or equivocal samples (2+) clustered below the 30% cellularity mark. Interestingly, the only sample with less than a 50% cellularity score was the sample with high intra-sample heterogeneity shown in Figure 2. In contrast, HercepTest-positive

(3+) amplified samples were widely scattered, with almost half of the cases showing ≤50% cellularity (one quarter showing <30%). The performances of HercepTest and VENTANA 4B5 methods were calculated considering SISH as the reference gold standard. Both immunohistochemistry tests showed excellent accuracy with ROC analysis, with VENTANA 4B5 marginally more accurate (VENTANA 4B5: area under curve: 0.98; 95% confidence interval (CI), 0.95–1.0; HercepTest: area under curve: 0.95; 95% CI, 0.85–1.0; P = 0.63; Supplementary Figure S3). Because of the limited number of cases, it was not possible to calculate the best cutoff value within separate immunohistochemistry scores. Therefore, we conservatively combined equivocal (2+) and positive (3+) staining samples and used two different empirical cutoffs for percentages of immunoreactive cells (≥10% or ≥50%). VENTANA 4B5 best performances are observed with a 50% cellularity cutoff showing 96.7% accuracy, 100% sensitivity, and 94.1% specificity (Table 2). At the same cutoff, HercepTest was more specific (100%) but considerably less sensitive (71.4%). VENTANA 4B5 always slightly outperformed HercepTest, with the latter's best performance, observed at a 10% cellularity cutoff, showing 90% accuracy, 92.9% sensitivity, and 87.5% specificity. False negatives were only present when HercepTest was used. False positive rates with VENTANA 4B5 were 30 and 3.3%, with a cellularity cutoff of 10% and 50%, respectively. The Panel of Consensus Pathologists also evaluated SISH performance in comparison to FISH by blind reading 30 paired slides randomly selected from both the positive/equivocal (N = 23) and the negative (N = 7) immunohistochemistry samples. Concordance was 100% (data not shown) with example patterns shown in Figure 1b. At the end of the review process, the Consensus Panel of Pathologists formulated a set of specific criteria for determining *ERBB2* positivity in colorectal cancer, referred to as HERACLES Diagnostic Criteria. The assessment of *ERBB2* positivity in colorectal cancer according to these criteria is a two-tier process whereby locally assessed samples, if negative (0/1+) or equivocal (2+) in <50% of cells, are excluded from further testing, while centralized immunohistochemistry and *in situ* hybridization re-testing is carried out according to specific staining intensities and cellularity cutoffs as reported in Table 3.

Clinical Validation Cohort (Step 2)

We screened 830 *KRAS* wild-type colorectal cancer patients according to HERACLES Diagnostic Criteria. *ERBB2* overexpression (2+/3+) and amplification prevalence rates are reported in Table 4 together with the summarized archival results for comparison. The prevalence of overexpressed cases significantly decreased from the archival to the screening cohort, from 13.7 to 8.4%, respectively (P = 0.02). On the contrary in both series the prevalence of

Table 2 Performances of two immunohistochemistry methods against *in situ* hybridization as the 'gold standard' for detection of *ERBB2* gene amplification

Collection	Test	Test positive if	Reference test	Number of patients					Test performances				
				True positive	*True negative*	*False positive*	*False negative*	*Total*	*Accuracy*	*Sensitivity*	*Specificity*	*Positive predictive value*	*Negative predictive value*
Archival CPP	HercepTest	2+ 3+ ≥10%	SISH	13	14	2	1	30	90.0%	92.9%	87.5%	86.7%	93.3%
Archival CPP	Ventana 4B5	2+ 3+ ≥10%	SISH	14	7	9	0	30	70.0%	100.0%	43.8%	60.9%	100.0%
Archival CPP	HercepTest	2+ 3+ ≥50%	SISH	10	16	0	4	30	86.7%	71.4%	100.0%	100.0%	80.0%
Archival CPP	Ventana 4B5	2+ 3+ ≥50%	SISH	13	16	1	0	30	96.7%	100.0%	94.1%	92.9%	100.0%

Abbreviations: CPP, Consensus of Pathologists; SISH, silver *in situ* hybridization.

Table 3 Consensus panel recommendations on *ERBB2* scoring for colorectal cancer (HERACLES Diagnostic Criteria)

Immunohistochemistry staining[a] at local pathology	Immunohistochemistry expected pattern	Immunohistochemistry classification	Referral to central pathology	Eligibility to HERACLES trial
No staining (0)	—	Negative	No	Not eligible
Faint staining (1+) any cellularity	Segmental or granular	Negative	No	Not eligible
Moderate (2+) in <50% cells	Any	Negative	No	Not eligible
Moderate (2+) in ≥50% of cells	Circumferential, basolateral or lateral	Equivocal	Mandatory: re-test immunohistochemistry. If >50% cellularity confirmed, proceed with *in situ* hybridization	If amplified[b], eligible
Intense (3+) in ≤10% cells	Circumferential, basolateral or lateral	Negative	No	Not eligible
Intense (3+) in >10% <50% of cells	Circumferential, basolateral or lateral	Positive	Mandatory: re-test immunohistochemistry. If >10% cellularity confirmed, proceed with *in situ* hybridization	If amplified[b], eligible
Intense (3+) in ≥50% of cells	Circumferential, basolateral or lateral	Positive	Confirmatory immunohistochemistry re-test. *In situ* hybridization not mandatory but recommended for research purposes	Eligible

[a]Recommendations relative to the use of VENTANA 4B5 immunohistochemistry assay and either FISH or SISH. [b]*ERBB2:CEP17* ratio ≥ 2 in ≥ 50% of cells.

Table 4 Prevalence of *ERBB2* expression and amplification in *KRAS* exon 2 WT patients in the archival and screening cohorts

Cohort	Patients	Immunohistochemistry 0			Immunohistochemistry 1+			Immunohistochemistry 2+				Immunohistochemistry 3+				Total overexpressed 2+/3+[a]		Total amplified[b]	
		N	%	Amplified[c]	N	%	Amplified[c]	N	%	Amplified[c]	% over 2+[a]	N	%	Amplified[c]	% over 3+	N	%	N	%
Archival[d]	256	131	51.2%	0	90	35.2%	0	23	9.0%	1	4.3%	12	4.7%	12	100.0%	35	13.7%	13	5.1%
Screening[d]	830	610	73.5%	—[e]	150	18.1%	—[e]	37	4.5%	10	27.0%	33	4.0%	33	100.0%	70	8.4%	43	5.2%
Total	1086	741	68.2%	NA	240	22.1%	NA	60	5.5%	11	18.3%	45	4.1%	45	100.0%	105	9.7%	56	5.2%

[a]Archival *vs* screening $\chi^2 = 4.94$; $P = 0.02$. [b]Archival *vs* screening = NS. [c]Archival: SISH. Screening: FISH. [d]Archival: scoring values from VENTANA 4B5 or HercepTest, whichever scored highest. Screening: Ventana 4B5 values. [e]According to HERACLES Diagnostic Criteria, ISH was not performed on 0/1+ samples.

confirmed *ERBB2* amplification was 5% (5.1 *vs* 5.2%). Interestingly all immunohistochemistry positive (3+) cases were amplified in both cohorts, while the percentage of amplified equivocal (2+) tumors increased from 4.3 to 27% from the archival to the screening cohort, respectively. VENTANA 4B5 test performances in the screening compared with the archival data sets using the ≥50% cellularity cutoff were identical for accuracy (96.7%), sensitivity (100%), and negative predictive value (100%). Specificity was higher in the screening data set than in the test data set (96.6 *vs* 94.1%, respectively), while, inversely, the positive predictive value decreased from 93 to 61% (Supplementary Table S2).

Discussion

While exploring *ERBB2* as an actionable therapeutic target in colorectal cancer in the HERACLES trial, by assessing the activity of dual ERBB2 inhibition with trastuzumab and lapatinib, we investigated the prevalence of *ERBB2* overexpression and amplification in more than a thousand prevalently metastatic colorectal cancer patients. The aim of the study was threefold.

First, a panel of *ERBB2* expert pathologists established that breast and gastric criteria for *ERBB2* positivity determination could also be made suitable to score *ERBB2* accurately and reproducibly in colorectal cancer. Technically, the Consensus of Pathologists Panel selected VENTANA 4B5 over HercepTest for protein expression determination because of lack of false negatives with the former. On the other hand, VENTANA 4B5, which recognizes both ERBB2 and ERBB4, had a higher false positive rate than the HercepTest, which uses a ERBB2-specific polyclonal antibody. This could be related to a cross-reaction of VENTANA 4B5 to ERBB4, or to operator-dependent factors resulting from the stronger retrieval processing procedure and the darker chromogen of the method. Given the relatively low prevalence of *ERBB2*-positive cases in colorectal cancer, the panel chose sensitivity over specificity to maximize potential *ERBB2*-positive patient's identification. The panel, recognizing that SISH analysis is more adaptable to processing of large sample batches (as in the archival collection) while FISH is more suitable for one-by-one screening, and because of the nearly perfect concordance between these two methods, also selected FISH analysis to determine gene amplification in the clinical validation cohort.

Second, we defined a *ERBB2* diagnostic algorithm, referred to as HERACLES Diagnostic Criteria. It focused on *ERBB2* amplification, as preclinical data supporting the HERACLES trial suggested amplification and not only overexpression as the predictive marker for response to anti-ERBB2 treatment.[10] Within this algorithm, to minimize false positives owing to unspecific staining, we disregarded cytoplasmic ERBB2 expression,[19,24] as in breast cancer

only membrane-bound ERBB2 expression is associated with ERBB2 gene amplification.[19] As in breast cancer, in our colorectal cancer series, ERBB2 overexpression associated with gene amplification was also already clearly membrane bound at low magnification. Interestingly, normal colon mucosa was generally stained with higher intensity than both the normal ductal epithelium and the normal gastric epithelium (Supplementary Figure S2). In breast and gastric cancer, the background staining noise can lead to reading biases, ultimately impinging on the concordance between ERBB2 immunohistochemistry and in situ hybridization.[25,26] The presence of a moderately intense normal background might thus have contributed to the scoring of the false negative case observed also with the HercepTest in our test cohort. Normalization protocols to deal with this issue in breast cancer have been both suggested[25] and highly criticized.[27] No formal normalization protocol was used in the present study, but the normal mucosa reactivity was a factor constantly considered by the pathologist during the scoring process, and the main reason why centralized testing has been recommended.

Third, to have a better estimate of the number of patients needing to be screened in order to achieve the intended sample size, we determined the prevalence of ERBB2 amplification in a representative population from the catchment area of the planned trial. In our study, the prevalence rate for ERBB2 amplification in KRAS wild-type samples was almost identical, i.e., 5.1 vs 5.2% in the archival test and in the clinical screening cohorts, respectively, thus prospectively validating the Heracles Diagnostic Criteria. ERBB2 overexpression rates were also almost identical for positive (3+) cases (4.7 vs 4.0%), whereas equivocal scored (2+) cases halved from the test to the validation cohort (9.0 vs 4.5%, respectively), suggesting a learning curve by local and central pathologists.

ERBB2 overexpression and amplification rates in colorectal cancer range widely owing to differences in technical approaches, antibodies, scoring protocols, cellular localization, and cellularity cutoffs.[6–9,12–20]

The present study is unique because the only study with an equally large sample size, albeit without clinical validation, reported much lower rates of both ERBB2 expression (2.7%) and positive amplification (1.6%).[8] This difference very likely resides in the fact that while our findings have been obtained on a primarily metastatic population, Heppner et al.[8] tested only primary tumors using a different staining antibody (SP3) and a different in situ hybridization method (CISH), and employed gastric cancer criteria to test and score the samples. Interestingly, the characteristics of ERBB2 positivity in our colorectal cancer samples are more similar to those of breast cancer than to gastric cancer.[3,22,28] ERBB2 staining in gastric cancer is not directly correlated to ERBB2 amplification, with up to 30% of

ERBB2 amplification-positive cases showing only focal staining or diffuse staining in < 30% of tumor cells.[24] On the contrary, ERBB2 protein expression and gene amplification in our colorectal cancer samples tallied quite accurately. None of the ERBB2 expression-negative colorectal cancer was amplified, whereas all immunohistochemistry-positive colorectal cancers were, leaving only the equivocal staining tumors in a grey zone. In addition, the cellularity of ERBB2 amplification is quite homogeneous with all positive cases displaying the amplification in >50% of cells. Another key point is intra-sample heterogeneity. In gastric cancer, areas strongly expressing ERBB2 at low magnification are often intermixed with areas displaying a much weaker membranous signal, and are only discernible at medium or high magnification corresponding to superimposable areas of high and no or patchy amplification.[21] A similar pattern was only observed in 1 out of the 53 amplified cases in our colorectal cancer series, precisely in the one case showing also high heterogeneity between primary tumor and liver metastases samples. On the other hand, ERBB2 membrane distribution in equivocal staining colorectal cancer samples is preferentially basolateral (U shaped) as in gastric[21] but not in breast cancer.[4,28] However, >70% of colorectacl cancer equivocal (2+) cases are not amplified and those that are amplified tend to show a slightly more patchwork-type intensity. Variability in the interpretation of immunohistochemistry staining and in situ hybridization results is to be expected in colorectal cancer, as testified by the still ongoing controversies on ERBB2 assessment in breast and gastric cancers, despite the continuous evolution of scoring guidelines.[21,23,29,30]

The crux of the matter, however, to ensure accurate identification of patients that might benefit from ERBB2-targeted therapy, across all tumor types, is whether the level of ERBB2 expression tallies or not with the level of ERBB2 amplification. In breast cancer, the level of ERBB2 protein is a known predictive factor of response to trastuzumab in metastatic disease[3] and of complete pathological response to the neoadjuvant therapy combination of lapatinib and trastuzumab.[31] Surprisingly, however, the level of amplification has been proven so far to be truly outcome predictive only for neoadjuvant therapy.[32–34] With gastric cancer in the ToGA trial, a trastuzumab-based regimen was found more effective in patients with tumors highly expressing the ERBB2 protein.[4] A positive correlation between levels of gene amplification and overall survival after trastuzumab-based therapy has also been recently established.[28] Early results from the HERACLES trial suggest that in colorectal cancer ERBB2 is a positive predictive marker for anti-ERBB2 targeted therapy.[35]

In conclusion, in a series of 1086 colorectal cancers, we optimized a diagnostic decisional algorithm referred to as HERACLES Diagnostic Criteria for ERBB2 protein overexpression and documented that ERBB2 amplification occurs in 5% of the population.

This study is the first clinical step toward a paradigmatic precision medicine path, initiated by functionally validating *ERBB2* as a therapeutic target in patient-derived xenografts from anti-EGFR therapy-resistant, *KRAS* wild-type metastatic colorectal cancer patients. Data from this study have provided important information on the stratification criteria for the HERACLES trial, a phase II clinical trial testing the combination of trastuzumab and lapatinib in metastatic colorectal cancer patients resistant to cetuximab or panitumumab.

Acknowledgments

Disclosure/conflict of interest

Andrea Sartore-Bianchi has received Honoraria by Amgen, Roche, and Bayer and has received payment for consulting or advisory roles by Amgen; for participation in Speaker's Bureaux and for travel, accommodation or expenses by Amgen, Roche, and Bayer. Frédérique Penaullt-Llorca has received Honoraria and has received payment for consulting or advisory roles, travel, accommodation or expenses by Roche and GlaxoSmithKline. Her institution has received research funding from Roche. Giuseppe Viale has received Honoraria by Roche, Novartis and payment for consulting or advisory roles by Roche, Dako, and Novartis. Sara Lonardi has received payment for consulting or advisory roles by Bayer and for participation in a Speaker's Bureau by Roche and received research funding from Sanofi-Aventis. Vittorina Zagonel has received payment for consulting or advisory roles by Bayer and Roche and for participation in Speaker's Bureaux by Roche, Amgen, Bayer, Merck Serono, Sanofi-Aventis, and Celgene. Johannes Noe is an employee of Hoffmann-La Roche. Fortunato Ciardiello has received payment for consulting or advisory roles by Bayer, Merck Serono, and Roche and has received research funding from Roche. Carmine Pinto's institution has received research funding from Merck Serono. Claudio Graiff has received payment for travel, accommodation or expenses by Pierre Fabre Pharma. Giuseppe Aprile has received payment for participation in Speaker's Bureaux by Amgen, Roche, and Lilly. Fotios Loupakis has received payment for consulting or advisory roles by Amgen and Sanofi-Aventis and for participation in Speaker's Bureaux by Roche, Sanofi-Aventis, Bayer, and Amgen and has received research funding from Roche and Merck Serono. Giuseppe Tonini has received Honoraria by Amgen, Roche, and Sanofi Aventis has received payment for consulting or advisory roles by Amgen, Roche and Sanofi Aventis and for participation in Speaker's Bureaux by Amgen, Roche, and Sanofi Aventis. Salvatore Siena has received payment for consulting or advisory roles by Amgen, Roche, Sanofi Aventis, Bayer, and Celgene. The other authors declare no conflict of interest.

References

1 Gutierrez C, Schiff R. HER2: biology, detection, and clinical implications. Arch Pathol Lab Med 2011;135: 55–62.

2 Arteaga CL, Engelman JA. ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics. Cancer Cell 2014;25:282–303.

3 Slamon DJ, Leyland-Jones B, Shak S et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001;344:783–792.

4 Bang YJ, 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. Lancet 2010;376:687–697.

5 Network CGA. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330–337.

6 Song Z, Deng Y, Zhuang K et al. Immunohistochemical results of HER2/neu protein expression assessed by rabbit monoclonal antibodies SP3 and 4B5 in colorectal carcinomas. Int J Clin Exp Pathol 2014;7:4454–4460.

7 Seo AN, Kwak Y, Kim DW et al. HER2 status in colorectal cancer: its clinical significance and the relationship between HER2 gene amplification and expression. PLoS One 2014;9:e98528.

8 Heppner BI, Behrens HM, Balschun K et al. HER2/neu testing in primary colorectal carcinoma. Br J Cancer 2014;111:1977–1984.

9 Park DI, Kang MS, Oh SJ et al. HER-2/neu overexpression is an independent prognostic factor in colorectal cancer. Int J Colorectal Dis 2007;22:491–497.

10 Bertotti A, Migliardi G, Galimi F et al. A molecularly annotated platform of patient-derived xenografts ("xenopatients") identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. Cancer Discov 2011;1:508–523.

11 Oncol JC (ed). Dual anti-HER2 treatment of patients with HER2-positive metastatic colorectal cancer: The HERACLES trial (HER2 Amplification for Colo-rectaL Cancer Enhanced Stratification). *Proceedings of the American Society Clinical Oncology Annual Meeting*; 31 May–4 June 2013. American Society of Clinical Oncology.

12 Tu J, Yu Y, Liu W *et al.* Significance of human epidermal growth factor receptor 2 expression in colorectal cancer. Exp Ther Med 2015;9:17–24.

13 Ramanathan RK, Hwang JJ, Zamboni WC *et al.* Low overexpression of HER-2/neu in advanced colorectal cancer limits the usefulness of trastuzumab (Herceptin) and irinotecan as therapy. A phase II trial. Cancer Invest 2004;22:858–865.

14 Ooi A, Takehana T, Li X *et al.* Protein overexpression and gene amplification of HER-2 and EGFR in color-ectal cancers: an immunohistochemical and fluores-cent in situ hybridization study. Mod Pathol 2004;17:895–904.

15 Pavlakis K, Kountourakis P, Stathopoulos E *et al.* Her-2 protein expression, cellular localization, and gene amplification in colorectal carcinoma. Appl Immunohistochem Mol Morphol 2007;15:441–445.

16 Kavanagh DO, Chambers G, O'Grady L *et al.* Is over-expression of HER-2 a predictor of prognosis in colo-rectal cancer? BMC Cancer 2009;9:1.

17 Marx AH, Burandt EC, Choschzick M *et al.* Hetero-genous high-level HER-2 amplification in a small subset of colorectal cancers. Hum Pathol 2010;41:1577–1585.

18 Kruszewski WJ, Rzepko R, Ciesielski M *et al.* Expres-sion of HER2 in colorectal cancer does not correlate with prognosis. Dis Markers 2010;29:207–212.

19 Blok EJ, Kuppen PJ, van Leeuwen JE *et al.* Cytoplasmic overexpression of HER2: a key factor in colorectal cancer. Clin Med Insights Oncol 2013;7:41–51.

20 Herreros-Villanueva M, Rodrigo M, Claver M *et al.* KRAS, BRAF, EGFR and HER2 gene status in a Spanish population of colorectal cancer. Mol Biol Rep 2011;38:1315–1320.

21 Hofmann M, Stoss O, Shi D *et al.* Assessment of a HER2 scoring system for gastric cancer: results from a validation study. Histopathology 2008;52:797–805.

22 Rüschoff J, Dietel M, Baretton G *et al.* HER2 diagnostics in gastric cancer-guideline validation and development of standardized immunohistochemical testing. Virch-ows Arch 2010;457:299–307.

23 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. J Clin Oncol 2007;25:118–145.

24 Ross JS, McKenna BJ. The HER-2/neu oncogene in tumors of the gastrointestinal tract. Cancer Invest 2001;19:554–568.

25 Gown AM, Goldstein LC, Barry TS *et al.* High concor-dance between immunohistochemistry and fluores-cence in situ hybridization testing for HER2 status in breast cancer requires a normalized IHC scoring system. Mod Pathol 2008;21:1271–1277.

26 Jeung J, Patel R, Vila L *et al.* Quantitation of HER2/neu expression in primary gastroesophageal adenocarcino-mas using conventional light microscopy and quanti-tative image analysis. Arch Pathol Lab Med 2012;136:610–617.

27 Hanna WM, Hammond E, Taylor CR *et al.* High concor-dance betweenimmunohistochemistry and fluorescence in situ hybridization testing for HER2 status inbreast cancer requires a normalized IHC scoring system. Mod Pathol 2008;21:1278–1280. (author reply 80–81).

28 Gomez-Martin C, Plaza JC, Pazo-Cid R *et al.* Level of HER2 gene amplification predicts response and overall survival in HER2-positive advanced gastric cancer treated with trastuzumab. J Clin Oncol 2013;31:4445–4452.

29 Wolff AC, Hammond ME, Hicks DG *et al.* Recommen-dations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol 2013;31:3997–4013.

30 Wolff AC, Hammond ME, Hicks DG *et al.* Recommen-dations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/college of American Pathologists clinical practice guideline update. Arch Pathol Lab Med 2014;138:241–256.

31 Scaltriti M, Nuciforo P, Bradbury I *et al.* High HER2 expression correlates with response to the combination of lapatinib and trastuzumab. Clin Cancer Res 2015;21:569–576.

32 Guiu S, Gauthier M, Coudert B *et al.* Pathological complete response and survival according to the level of HER-2 amplification after trastuzumab-based neoad-juvant therapy for breast cancer. Br J Cancer 2010;103:1335–1342.

33 Gullo G, Bettio D, Torri V *et al.* Level of HER2/neu gene amplification as a predictive factor of response to trastuzumab-based therapy in patients with HER2-positive metastatic breast cancer. Invest New Drugs 2009;27:179–183.

34 Piccart-Gebhart MJ, Procter M, Leyland-Jones B *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med 2005;353:1659–1672.

35 Siena S, Sartore-Bianchi A, Lonardi S *et al.* Trastuzu-mab and lapatinib in HER2 amplified metastatic color-ectal cancer patients (mCRC): the HERACLES trial. J Clin Oncol 2015;33:Suppl 3508. (oral presentation at ASCO Annual Meeting, Chicago, IL, USA, 28 May–2 June 2015).