

Value of staining intensity in the interpretation of immunohistochemistry for tumor markers in colorectal cancer

Inti Zlobec · Luigi Terracciano · Jeremy R. Jass ·
Alessandro Lugli

Received: 9 May 2007 / Accepted: 27 June 2007 / Published online: 3 August 2007
© Springer-Verlag 2007

Abstract The purpose of this study was to determine whether staining intensity in conjunction with the percentage of positive tumor cells should be used as an indicator of protein expression detected by immunohistochemistry. A tissue microarray of 1,197 colorectal cancers was immunostained for p53, Her2/neu, epidermal growth factor receptor (EGFR), adenomatous polyposis coli (APC), and β -catenin. Immunoreactivity was described by the percentage of positive tumor cells (percent positivity) and by the staining intensity (weak, moderate, strong). The interobserver reproducibility of both was evaluated by two pathologists. The association of T stage, N stage, tumor grade, vascular invasion, and survival with percent positivity, staining intensity, and the combination of both was assessed. In univariate analysis, protein expression assessed by percent positivity resulted in 11 significant associations between the proteins and clinico-pathological features. Eight of these 11 were also demonstrated using only the degree of staining intensity. However, more than half of the associations identified by percent positivity alone were lost when staining intensity was also analyzed in combination with the percentage

of positive tumor cells. A scoring method based on percent positivity, rather than on staining intensity, for p53, Her2/neu, EGFR, APC, and β -catenin is reproducible and appears to be sufficient for establishing associations of the selected tumor markers with most clinico-pathological features.

Keywords Colorectal cancer · Immunohistochemistry · Scoring system · Staining intensity

Introduction

Diagnosis and research in colorectal cancer (CRC) routinely relies on the interpretation of protein expression detected by means of immunohistochemistry (IHC). The lack of standardized IHC scoring systems has led to a variety of unvalidated methods used to assess immunoreactivity. Scoring systems for tumor markers in CRC are usually based on a measure of the proportion of positive tumor cells and are often combined with a degree of staining intensity [3, 9, 12, 29]. It is recognized that the interpretation of staining intensity is not only highly subjective but may be affected by storage time, variation in protocols, and fixation procedures [2, 13]. Despite these concerns, staining intensity has become an integral component of many IHC scoring methods for tumor markers in CRC [7, 8, 10, 12, 23, 25–27].

We have recently shown that a descriptive, semiquantitative scoring system for IHC based on the percentage of immunoreactive tumor cells (percent positivity) provides a more complete assessment of the predictive and prognostic value of several tumor markers in CRC when compared to an evaluation system using “negative/positive” [16–19]. Additionally, we have studied the interobserver variability of this semiquantitative method among pathologists for several proteins namely, epidermal growth factor receptor

I. Zlobec (✉) · L. Terracciano · A. Lugli
Institute of Pathology, University Hospital of Basel,
Schönbeinstrasse 40,
Basel 4031, Switzerland
e-mail: izlobec@uhbs.ch

J. R. Jass
Department of Cellular Pathology, St. Mark's Hospital,
London, UK

J. R. Jass
Imperial College,
London, UK

(EGFR), vascular endothelial growth factor, p53, Bcl-2, and apoptotic protease-activating factor 1 (APAF-1), in tumor biopsies and tissue microarray (TMA) punches and have shown that it is reproducible [30, 31, 34].

One of the most important advantages of a semiquantitative scoring method is that it allows the investigator to establish more biologically or clinically relevant cutoff scores for positivity for the protein and the outcome under study rather than relying on an often arbitrary threshold value, such as 10%, to describe a tumor as “positive.” Such a method to ascertain cutoff scores has been recently proposed using receiver operating characteristic (ROC) curve analysis and has been applied, along with several other tumor markers, to EGFR and APAF-1 [32–34].

The purpose of this study was to determine whether staining intensity in conjunction with percent positivity should be used as an indicator of protein expression detected by IHC. The associations of p53-, Her2/neu-, EGFR-, adenomatous polyposis coli (APC)-, and β -catenin-staining intensity with a range of clinico-pathological features, notably T stage, N stage, tumor grade, vascular invasion, and survival, were evaluated in 1,197 mismatch repair (MMR)-proficient CRC.

Materials and methods

Tissue microarray construction

A TMA of 1,420 unselected, nonconsecutive CRCs was constructed [24]. Briefly, formalin-fixed, paraffin-embedded tissue blocks of CRC resections were obtained. One tissue cylinder with a diameter of 0.6 mm was punched from morphologically representative tissue areas of each donor tissue block and brought into one recipient paraffin block (3×2.5 cm) using a homemade semiautomated tissue arrayer. The resulting TMA set comprised three slides.

Clinico-pathological data

The clinico-pathological data for all patients included T stage (T1, T2, T3, and T4), N stage (N0, N1, and N2), tumor grade (G1, G2, and G3), vascular invasion (presence or absence), and survival time [19].

Immunohistochemistry

Four-micrometer sections of TMA blocks were transferred to an adhesive-coated slide system (Instrumedics, Hackensack, NJ). Standard indirect immunoperoxidase procedures were used for IHC (ABC-Elite, Vector Laboratories, Burlingame, CA). One thousand four hundred and twenty CRCs were immunostained for mutL homolog (MLH)1 (clone MLH-1;

dilution 1:100; BD Biosciences Pharmingen, San Jose, CA), MSH2 (clone MSH-2; dilution 1:200; BD Biosciences Pharmingen, San Jose, CA), and MSH6 (clone 44; dilution 1:400, Transduction Laboratories). After dewaxing and rehydration in deionized H₂O, sections were subjected to heat antigen retrieval in a microwave oven (1,200 W, 15 min) in 0.001 mol/L ethylenediamine tetraacetic acid at pH 8.0 for MLH1 and MSH2 and in 0.01 mol/L citrate buffer (pH 7.0) for MSH6. Endogenous peroxidase activity was blocked using 0.5% H₂O₂. The sections were incubated with 10% normal goat serum (Dako Cytomation, Mississauga, Canada) for 20 min and incubated with a primary antibody at room temperature (Her2/neu clone PN2A, DAKO, Denmark; p53 clone DO-7, 1:100; DAKO; clone β -catenin-1, dilution 1:200; Dako Cytomation, APC clone C20, dilution 1:50; Santa Cruz, CA). Subsequently, sections were incubated with secondary antibody (K4005, EnVision+ System-HRP (AEC); Dako Cytomation) for 30 min at room temperature. For visualization of the antigen, the sections were immersed in 3-amino-9-ethylcarbazole+substrate-chromogen (K4005, EnVision+ System-HRP (AEC); Dako Cytomation) for 30 min and counterstained with Gill's hematoxylin. IHC for EGFR was performed using an automated stainer (EGFR clone 3C6, 3 mg/mL; Ventana Medical Systems, Tucson, USA) according to the manufacturer's instructions. TMA slides for each protein were stained on the same day under identical conditions.

MMR results

The 1,420 CRCs were stratified according to MMR status: (1) MMR-proficient tumors expressing MLH1, MSH2, and MSH6, (2) MLH1-negative tumors, and (3) presumed hereditary nonpolyposis colorectal cancer (HNPCC) cases demonstrating loss of MSH2 and/or MSH6 at any age or loss of MLH1 at age less than 55 years [11]. In population-based studies, the mean age of diagnosis of Lynch syndrome is around 55 years. By contrast, sporadic high-level microsatellite instability (MSI-H; MMR-deficient) CRC is far more age-related with a mean age of onset of around 75 years and few cases occurring younger than 60 years. A cutoff of 55 years was set as a reasonable compromise for distinguishing CRCs with loss of MLH1 into likely HNPCC syndrome vs likely sporadic. These immunohistochemical groupings showed a good fit with the known clinicopathological features associated with these subsets of CRC. Particularly, the MLH1-negative group was associated with advanced age, predilection for women and the proximal colon, large tumor size, and poor differentiation. The presumed HNPCC group was young and showed no gender difference, and there was a predilection for the proximal colon as compared with the MMR-proficient group. While it is possible that a small proportion of presumed sporadic MSI-H and HNPCC

cases were incorrectly assigned, the overall findings are likely to be valid in view of the large numbers of samples and the good fit with clinico-pathological features. Only MMR-proficient tumors were included in the study to ensure a homogeneous sample of tumors ($N=1,197$, 84.4%).

IHC evaluation

Immunoreactivity was evaluated in all 1,420 punches by one experienced pathologist. TMA punches with insufficient tissue or tumor for analysis were excluded from the study. Protein expression was scored in the nucleus for p53 and β -catenin and in the cytoplasm for APC. EGFR and Her2/neu positivity were scored in both cell membrane and cytoplasm. Immunoreactivity was assigned a score based on the proportion of positive tumor cells over total tumor cells (percent positivity) ranging from 0 to 100%. Staining intensity was evaluated as 0=negative, 1=weak, 2=moderate, and 3=strong. If the staining intensity was heterogeneous, then scoring was based on the greatest degree of intensity. MLH1, MSH2, and MSH6 were scored in the nucleus as negative (0%) and as positive (>0%).

Statistical analysis

Interobserver reproducibility of scoring percent positivity and staining intensity

To determine the interobserver reproducibility of percent positivity and staining intensity, a minimum of 100 CRC punches was evaluated by a second pathologist. The intraclass correlation coefficient (ICC) defined as the ratio of the between-subject variance over the between-subject+within-subject variances was used to determine the reliability of percent positivity for each protein. The ICC has previously been used to assess the agreement of IHC scores [14, 31]. An ICC of 0.7 or greater is considered sufficient to establish reproducibility [14]. The interobserver agreement of staining intensity (negative, weak, moderate, and strong) was determined using the kappa coefficient (κ) [15]. The overall κ coefficient measures the reliability of categorical data while taking into account the probability that both observers achieved the same scores by chance [1]. The weighted κ may be used as a measure of inter-rater agreement for ordinal variables and quantifies the relative difference between them. The greater the difference is between the scores, the lower the weighted κ . The interpretation of κ is commonly made as follows: values between 0.81 and 1.0 represent “almost perfect” agreement, 0.61 and 0.80 “substantial” agreement, 0.41 and 0.60 “moderate” agreement, 0.21 and 0.40 “fair” agreement, and 0 and 0.20 “slight” agreement [15].

Univariate analysis

Logistic regression analysis was used to determine the association of percent positivity or staining intensity with T stage (early=T1+T2/late=T3+T4), N stage (absence [N0] or presence [>N0] of lymph node involvement), tumor grade (low=G1+G2/high=G3), and vascular invasion (absence or presence). Survival analysis was carried out using Cox proportional hazards regression.

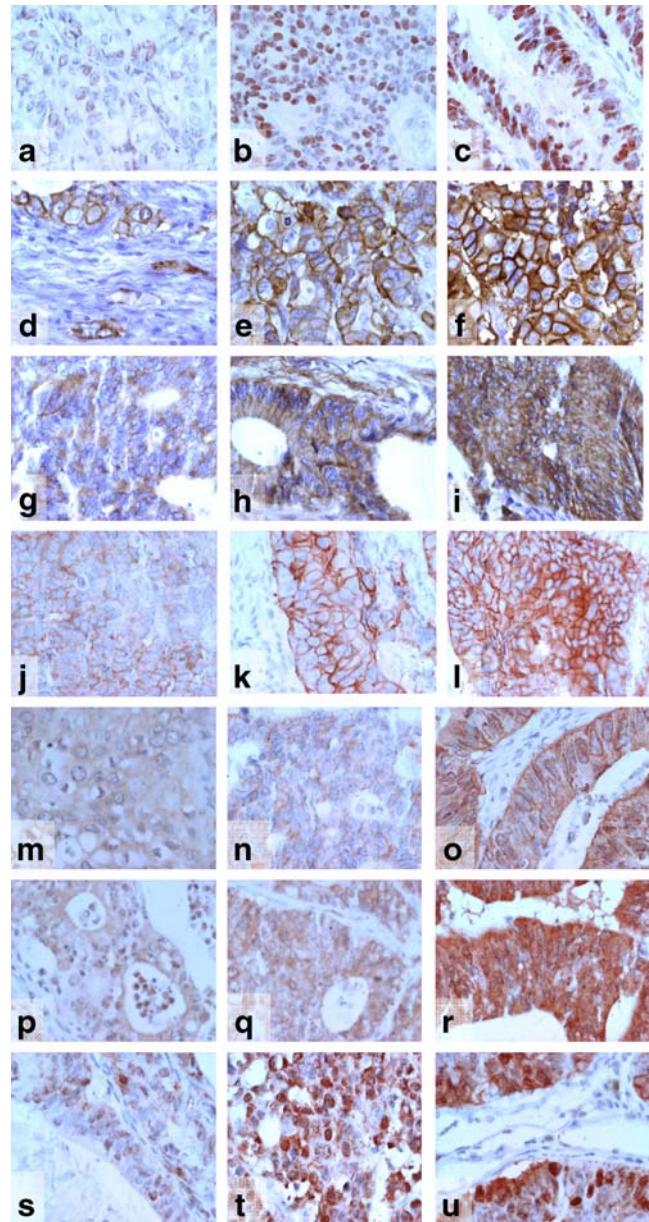


Fig 1 a–l Weak (left), moderate (center), and strong (right) staining intensity of nuclear p53 (40 \times ; a–c), membranous EGFR (40 \times ; d–f), cytoplasmic EGFR (40 \times ; g–i), and membranous Her2/neu (40 \times ; j–l). m–o Weak (left), moderate (center), and strong (right) staining intensity of cytoplasmic Her2/neu (40 \times ; m–o), cytoplasmic APC (40 \times ; p–r), and nuclear β -catenin (40 \times ; s–u)

Table 1 Interobserver agreement for percent positivity (percent scores) measured by the intraclass correlation coefficient (ICC) and for staining intensity measured by the overall and weighted kappa coefficient (κ)

Protein	Measurements of interobserver agreement		
	Percent scores	Staining intensity	
	ICC	Overall κ (95%CI)	Weighted κ (95%CI)
p53 (n)	0.91	0.2 (0.11–0.3)	0.5 (0.43–0.57)
APC (c)	0.85	0.41 (0.23–0.58)	0.53 (0.36–0.68)
β -Catenin (n)	0.78	0.34 (0.22–0.47)	0.52 (0.39–0.64)
Her2/neu (m)	0.71	0.53 (0.36–0.69)	0.67 (0.52–0.82)
Her2/neu (c)	0.68	0.57 (0.43–0.71)	0.62 (0.47–0.76)
EGFR (c)	0.54	0.12 (0.0–0.25)	0.26 (0.14–0.39)
EGFR (m)	0.48	0.11 (0.01–0.22)	0.32 (0.22–0.42)

n Nuclear, c cytoplasmic,
m membranous

Multivariate analysis

Staining intensity and percent positivity were entered into a multivariate logistic regression model for all binary outcomes, whereas multiple Cox proportional hazards regression was performed for survival analysis. Adjusted *P* values for percent positivity and staining intensity were obtained.

P values less than or equal to 0.05 indicate a significant association of percent positivity or staining intensity with the outcome. All analyses were carried out using SAS version 9.1 (The SAS Institute, Cary, NC).

Results

Interobserver reproducibility of percent positivity and staining intensity

The reproducibility of scores expressed as percent positivity was very strong for p53 (ICC=0.91). The ICCs for APC (ICC=0.85), β -catenin (ICC=0.78), and membranous Her2/neu (ICC=0.71) suggest excellent consistency of scores between observers. The interobserver agreement for cytoplasmic Her2/neu scores was only slightly lower than for its membranous counterpart (ICC=0.68). However, the reliability of membranous and cytoplasmic EGFR expression was only moderate to low. The same observers independently re-evaluated EGFR expression a second time by scoring the number of immunoreactive tumor cells without regard to

localization of staining (i.e., by scoring either membrane and/or cytoplasmic EGFR). The interobserver agreement was significantly increased (ICC=0.86) [33].

The agreement of staining intensity between observers using the overall κ coefficient was only moderate for APC (κ =0.41) and membranous and cytoplasmic Her2/neu (κ =0.53 and 0.57, respectively). The reproducibility of staining intensity for β -catenin was determined to be fair (κ =0.34) while that of membranous or cytoplasmic EGFR was poor (κ =0.11 and 0.12, respectively). Analyses with the more generous weighted κ did little to improve these findings (Fig. 1; Table 1).

Association of protein expression with clinico-pathological features

Univariate analysis of percent scores

The evaluation of percent positivity with the clinico-pathological features identified significant associations between p53 and T stage (P =0.007) and tumor grade (P =0.005), membranous EGFR expression and T stage (P =0.005), N stage (P =0.002), tumor grade (P =0.014), and survival (P <0.001), and cytoplasmic EGFR expression and survival (P =0.01). APC expression was correlated with T stage (P =0.023) and β -catenin expression with tumor grade (P =0.035), vascular invasion (P =0.008), and survival (P =0.004). There were no associations of Her2/neu with any of the clinico-pathological features (Table 2).

Table 2 Association of protein expressed as percent positivity with clinico-pathological features by univariate analysis (*P* value)

	T stage	N stage	Tumor grade	Vascular invasion	Survival
p53 (n)	0.007	0.365	0.005	0.358	0.401
EGFR (m)	0.005	0.002	0.014	0.128	<0.001
EGFR (c)	0.193	0.692	0.382	0.522	0.01
Her2/neu (m)	0.857	0.299	0.551	0.475	0.932
Her2/neu (c)	0.322	0.629	0.235	0.139	0.15
APC (c)	0.023	0.138	0.976	0.205	0.208
β -Catenin (n)	0.619	0.073	0.035	0.008	0.004

n Nuclear, c cytoplasmic,
m membranous

Table 3 Association of staining intensity with clinico-pathological features by univariate analysis (*P* value)

	T stage	N stage	Tumor grade	Vascular invasion	Survival
p53 (n)	0.026	0.127	0.017	0.991	0.61
EGFR (m)	0.005	0.009	0.049	0.192	<0.001
EGFR (c)	0.108	0.447	0.736	0.908	0.017
Her2/neu (m)	0.101	0.51	0.657	0.632	0.738
Her2/neu (c)	0.671	0.487	0.2	0.384	0.364
APC (c)	0.029	0.866	0.714	0.942	0.689
β -catenin (n)	0.424	0.541	<0.001	0.067	0.175

n Nuclear, *c* cytoplasmic, *m* membranous

Univariate analysis of staining intensity

All eight significant associations of staining intensity with the clinico-pathological features were previously established by percent positivity. The associations of membranous EGFR and tumor grade, β -catenin, vascular invasion, and survival were not present (Table 3).

Multivariate analysis of percent positivity and staining intensity

The combined analysis of percent positivity with staining intensity identified five associations between the proteins and clinico-pathological features that were previously found using only the percentage of positive cells. In addition, the remaining associations found to have statistical significance in univariate analysis of percent positivity (p53 with T stage, membranous EGFR with T stage, N stage, and tumor grade, APC with T stage, and β -catenin with tumor grade) were no longer observed in combination with the degree of intensity. In only 1 of the 35 analyses (2.8%) did staining intensity provide additional information about the association of the protein with the outcome (β -catenin with N stage; Table 4).

Discussion

The results of this study confirm that the evaluation of percent positivity for nuclear p53, cytoplasmic APC, nuclear β -catenin, and membranous and cytoplasmic Her2/neu expression is highly reproducible among pathologists. The assessment of EGFR expression resulted in strong interobserver agreement when cytoplasmic and/or membranous immunoreactivity were scored together rather than in their separate localizations. The intensity of staining was not reproducible for the proteins in this study.

In the univariate analysis, protein expression assessed as percent positivity resulted in 11 significant associations between the proteins and clinico-pathological features. Eight of these 11 were also demonstrated using only the degree of staining intensity. However, more than half of the associations identified by percent positivity alone were lost when staining intensity was also analyzed.

Scoring systems for tumor markers in CRC are typically based on some measure of the number of positive tumor cells and often combined with a degree of staining intensity. However, Atkins et al. [2] demonstrated using an anti-EGFR antibody in head and neck cancer, non-small cell lung carcinomas, and colorectal adenocarcinoma that the degree

Table 4 Association of IHC expressed as percent positivity and clinico-pathological features in multivariate analysis with staining intensity (adjusted *P* values)

	T stage	N stage	Tumor grade	Vascular invasion	Survival
p53 (n)	0.09 ^a	0.643	0.01 ^b	0.745	0.409
EGFR (m)	0.095 ^a	0.19 ^a	0.341 ^a	0.676	0.021 ^b
EGFR (c)	0.777	0.815	0.266	0.339	0.045 ^b
Her2/neu (m)	0.438	0.32	0.46	0.804	0.693
Her2/neu (c)	0.125	0.7	0.327	0.322	0.167
APC (c)	0.137 ^a	0.084	0.815	0.125	0.179
β -catenin (n)	0.446	0.025 ^c	0.062 ^a	0.036 ^b	0.005 ^b

n Nuclear, *c* cytoplasmic, *m* membranous

^a Loss of significance

^b Continued significance

^c Gain of significance with the addition of staining intensity

of staining intensity varied by tumor type, was partially influenced by the choice of fixatives, and was inversely correlated with storage time of the unstained tissue sections. These factors in addition to the variation in IHC protocols inevitably contribute to the subjective nature of staining intensity. Contradictory results from different reports on the same tumor markers may be partially explained by this subjective assessment of immunoreactivity [23, 26].

We have previously demonstrated that a descriptive, semiquantitative scoring system based on the percentage of positive tumor cells (percent positivity) is reproducible and has several advantages over standard scoring methods based on predetermined cutoff scores [17–19]. First, this scoring system allows a more thorough assessment of the predictive or prognostic significance of tumor markers by evaluating the entire range of protein expression levels (from 0 to 100%). Moreover, by quantifying protein expression at the outset, more biologically and clinically relevant cutoff scores for tumor positivity can be established by, for example, performing ROC curve analysis [34]. This method has been used to select cutoff scores for tumor markers macrophage stimulating factor 1, Raf-1 kinase inhibitor protein, receptor of hyaluronic acid-mediated motility (RHAMM), APAF-1, EGFR, as well as for several others involved in transforming growth factor β signaling in CRC [4, 20, 21, 32, 33]. Additionally, the correlations between various proteins can be assessed. We have recently shown using this scoring method that the percentage of pERK-positive tumor cells is strongly associated with increases in RHAMM expression supporting the hypothesis of a RHAMM–mitogen-activated protein kinase interaction in MMR-proficient CRC [17]. By percent scoring, we have also described how classification and regression tree methods could be used to select proteins playing a role in predicting rectal tumor response to preoperative radiotherapy [30]. Finally, this descriptive scoring method avoids an often complex and interpretative composite scoring system based on the intensity of staining. One such method includes a four-tier scoring of the intensity of staining (0, 1+, 2+, 3+) coupled to either the mean percentage of positive tumor cells or to a categorical measure of the percentage of positive tumor cells (for example, 1–10%, 10–50%, and >50) [3, 6, 9, 12, 29]. A graded scoring system has also been used where the percentage of positive tumor cells is categorized (0=no positivity, 1=1–25%, 2=25–50%, 3=>50%) and multiplied by the degree of intensity (0, 1, 2, 3) to obtain a score that is then dichotomized into “low” or “high” expression (low=score<6 and high=score \geq 6) [26]. Others have reported only the degree of staining intensity regardless of the proportion of immunoreactivity or considered only staining intensities of 2+ or 3+ as positive for protein expression [24].

The purpose of this study was not to evaluate the prognostic significance of several tumor markers in CRC

but rather to determine whether staining intensity is a useful indicator of immunoreactivity in colorectal tumors. In addition, the study focused on whether staining intensity provides independent information on the association of the protein with clinico-pathological features beyond that which can be obtained from the semiquantitative assessment of immunoreactivity. The markers included in this study are well established and/or of current interest as prognostic factors. They were selected to provide a range of subcellular localizations for scoring purposes (cytoplasm, cell membrane, nucleus) as well as representing both tumor suppressors (p53 and APC) and oncogenes (β -catenin, Her2/neu, and EGFR).

TMA technology allowed us to analyze more than 1,000 CRCs using only three slides. One tissue sample (0.6 mm) per tumor was obtained. Although it is argued that a single tissue core may not be representative of the whole tumor, results using one sample appear to approximate those from larger tissue sections as more samples are analyzed. In fact, even larger tissue sections may contain only a small fraction of the entire tumor mass (1/10,000) [24]. Goethals et al. [9] reported that four core biopsies are sufficient to account for tumor heterogeneity. Because the inclusion of several punches per tumor is not always possible, a larger series of tumors should compensate for tumor heterogeneity as was the case in this study. Several studies have shown well-established associations between molecular features and clinico-pathological endpoints in TMAs using only one spot per tumor [5, 22, 28]. Most importantly, evaluating a single tumor punch may lead to a more reliable analysis of interobserver agreement for both percent positivity and staining intensity, as precisely the same area of tumor is scored by each observer [24].

The results of our study suggest that staining intensity is not an independent measure of protein expression for the markers in this study. Additionally, the evaluation of immunoreactivity using a semiquantitative scoring method appears to be sufficient for establishing associations of the selected tumor markers with most clinico-pathological features.

References

1. Ansari-Lari M (2005) Letter to the editor: comparison between two test results, k statistic instead of simple overall agreement. *Vet Parasitol* 133:369–370
2. Atkins D, Reiffen KA, Tegtmeier CL, Winther H, Bonato MS, Storkel S (2004) Immunohistochemical detection of EGFR in paraffin-embedded tumor tissues: variation in staining intensity due to choice of fixative and storage time of tissue sections. *J Histochem Cytochem* 52:893–901
3. Azria D, Bibeau F, Barbier N, Zouhair A, Lemanski C, Rouanet P, Ychou M, Senesse P, Ozsahin M, Pelegrin A, Dubois JB, Thezenas S (2005) Prognostic impact of epidermal growth factor

- receptor (EGFR) expression on loco-regional recurrence after preoperative radiotherapy in rectal cancer. *BMC Cancer* 5:62
4. Baker K, Zlobec I, Tornillo L, Terracciano L, Jass JR, Lugli A (2007) Differential significance of tumour infiltrating lymphocytes in sporadic mismatch repair deficient versus proficient colorectal cancers: a potential role for dysregulation of the transforming growth factor-beta pathway. *Eur J Cancer* 43:624–631
 5. Barlund M, Forozan F, Kononen J, Bubendorf L, Chen Y, Bittner ML, Torhorst J, Haas P, Bucher C, Sauter G, Kallioniemi OP, Kallioniemi A (2000) Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. *J Natl Cancer Inst* 92:1252–1259
 6. Bibeau F, Boissiere-Michot F, Sabourin JC, Gourgou-Bourgade S, Radal M, Penault-Llorca F, Roचाix P, Arnould L, Bralet GP, Azria D, Ychou M (2006) Assessment of epidermal growth factor receptor (EGFR) expression in primary colorectal carcinomas and their related metastases on tissue sections and tissue microarray. *Virchows Arch* 449:281–287
 7. Bravou V, Klironomos G, Papadaki E, Taraviras S, Varakis J (2006) ILK over-expression in human colon cancer progression correlates with activation of beta-catenin, down-regulation of E-cadherin and activation of the Akt-FKHR pathway. *J Pathol* 208:91–99
 8. Garrity MM, Burgart LJ, Mahoney MR, Windschitl HE, Salim M, Wiesenfeld M, Krook JE, Michalak JC, Goldberg RM, O'Connell MJ, Furth AF, Sargent DJ, Murphy LM, Hill E, Riehle DL, Meyers CH, Witzig TE (2004) Prognostic value of proliferation, apoptosis, defective DNA mismatch repair, and p53 overexpression in patients with resected Dukes' B2 or C colon cancer: a North Central Cancer Treatment Group Study. *J Clin Oncol* 22:1572–1582
 9. Goethals L, Perneel C, Debucquoy A, De Schutter H, Borghys D, Ectors N, Geboes K, McBride WH, Haustermans KM (2006) A new approach to the validation of tissue microarrays. *J Pathol* 20: 607–614
 10. Goldstein NS, Armin M (2001) Epidermal growth factor receptor immunohistochemical reactivity in patients with American Joint Committee on Cancer Stage IV colon adenocarcinoma: implications for a standardized scoring system. *Cancer* 92:1331–1346
 11. Hampel H, Stephens JA, Pukkala E, Sankila R, Aaltonen LA, Mecklin JP, de la Chapelle A (2005) Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology* 129:415–421
 12. Italiano A, Saint-Paul MC, Caroli-Bosc FX, Francois E, Bourgeon A, Benchimol D, Gugenheim J, Michiels JF (2005) Epidermal growth factor receptor (EGFR) status in primary colorectal tumors correlates with EGFR expression in related metastatic sites: biological and clinical implications. *Ann Oncol* 16:1503–1507
 13. Italiano A, Vandenbos FB, Otto J, Mouroux J, Fontaine D, Marcy PY, Cardot N, Thyss A, Pedeutour F (2006) Comparison of the epidermal growth factor receptor gene and protein in primary non-small-cell-lung cancer and metastatic sites: implications for treatment with EGFR-inhibitors. *Ann Oncol* 17:981–985
 14. Kirkegaard T, Edwards J, Tovey S, McGlynn LM, Krishna SN, Mukherjee R, Tam L, Munro AF, Dunne B, Bartlett JM (2006) Observer variation in immunohistochemical analysis of protein expression, time for a change? *Histopathology* 48:787–794
 15. Landis J, Koch G (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
 16. Lugli A, Zlobec I, Baker K, Minoo P, Tornillo L, Terracciano L, Jass J (2007) Prognostic significance of mucins in colorectal cancer with different DNA mismatch-repair status. *J Clin Pathol* 60:534–539
 17. Lugli A, Zlobec I, Gunthert U, Minoo P, Baker K, Tornillo L, Terracciano L, Jass JR (2006) Overexpression of the receptor for hyaluronic acid mediated motility is an independent adverse prognostic factor in colorectal cancer. *Mod Pathol* 19:1302–1309
 18. Lugli A, Zlobec I, Minoo P, Baker K, Tornillo L, Terracciano L, Jass JR (2006) Prognostic significance of the wnt signaling pathway molecules APC, β -catenin and E-cadherin in colorectal cancer. *Histopathology* 50:454–464
 19. Lugli A, Zlobec I, Minoo P, Baker K, Tornillo L, Terracciano L, Jass JR (2006) Role of the mitogen-activated protein kinase and phosphoinositide 3-kinase/AKT pathways downstream molecules, phosphorylated extracellular signal-regulated kinase, and phosphorylated AKT in colorectal cancer-A tissue microarray-based approach. *Human Pathol* 37:1022–1031
 20. Minoo P, Zlobec I, Baker K, Tornillo L, Terracciano L, Jass JR, Lugli A (2006) Loss of Raf-1 kinase inhibitor protein (RKIP) expression is associated with tumor progression and metastasis in colorectal cancer. *Am J Clin Pathol* 127:820–827
 21. Minoo P, Zlobec I, Baker K, Tornillo L, Terracciano L, Jass JR, Lugli A (2007) Prognostic significance of mammalian sterile-20-like kinase 1 in colorectal cancer. *Mod Pathol* 20:331–338
 22. Moch H, Schraml P, Bubendorf L, Mirlacher M, Kononen J, Gasser T, Mihatsch MJ, Kallioniemi OP, Sauter G (1999) High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. *Am J Pathol* 154:981–986
 23. Resnick MB, Routhier J, Konkin T, Sabo E, Pricolo VE (2004) Epidermal growth factor receptor, c-MET, beta-catenin, and p53 expression as prognostic indicators in stage II colon cancer: a tissue microarray study. *Clin Cancer Res* 10:3069–3075
 24. Sauter G, Simon R, Hillan K (2003) Tissue microarrays in drug discovery. *Nat Rev Drug Discov* 2:962–972
 25. Schwandner O, Schiedeck TH, Bruch HP, Duchrow M, Windhoevel U (2000) p53 and Bcl-2 as significant predictors of recurrence and survival in rectal cancer. *Eur J Cancer* 36:348–356
 26. Spano JP, Lagorce C, Atlan D, Milano G, Domont J, Benamouzig R, Attar A (2005) Impact of EGFR expression on colorectal cancer patient prognosis and survival. *Ann Oncol* 16:102–108
 27. Takahashi Y, Bucana CD, Cleary KR, Ellis LM (1998) p53, vessel count, and vascular endothelial growth factor expression in human colon cancer. *Int J Cancer* 79:34–38
 28. Torhorst J, Bucher C, Kononen J, Haas P, Zuber M, Kochli OR, Mross F (2001) Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 159:2249–2256
 29. Umemura S, Itoh J, Itoh H, Serizawa A, Saito Y, Suzuki Y, Tokuda Y, Tajima T, Osamura RY (2004) Immunohistochemical evaluation of hormone receptors in breast cancer: which scoring system is suitable for highly sensitive procedures? *Appl Immunohistochem Mol Morphol* 12:8–13
 30. Zlobec I, Steele R, Nigam N, Compton CC (2005) A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. *Clin Cancer Res* 11:5440–5443
 31. Zlobec I, Steele R, Michel RP, Compton CC, Lugli A, Jass JR (2006) Scoring of p53, VEGF, Bcl-2 and APAF-1 immunohistochemistry and interobserver reliability in colorectal cancer. *Mod Pathol* 19:1236–1242
 32. Zlobec I, Minoo P, Baker K, Haegert D, Khetani K, Tornillo L, Terracciano L, Jass JR, Lugli A (2007) Loss of APAF-1 expression is associated with tumour progression and adverse prognosis in colorectal cancer. *Eur J Cancer* 43:1101–1107
 33. Zlobec I, Vuong T, Hayashi S, Haegert D, Tornillo L, Terracciano L, Lugli A, Jass J (2007) A simple and reproducible scoring system for EGFR in colorectal cancer: application to prognosis and prediction of response to preoperative brachytherapy. *Br J Cancer* 96:793–800
 34. Zlobec I, Steele R, Terracciano L, Jass JR, Lugli A (2007) Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer. *J Clin Pathol* (in press). DOI 10.1136/jcp.2006.044537