Journal of Surgical Oncology 2016;113:46–54

Ki67—An Unsuitable Marker of Gastric Cancer Prognosis Unmasks Intratumoral Heterogeneity

CHRISTINE BÖGER, MD, HANS-MICHAEL BEHRENS, MSc, AND CHRISTOPH RÖCKEN, PhD*
Department of Pathology, Christian-Albrechts-University, Kiel, Germany

Background and Objectives: Due to contradictious findings of previous studies regarding Ki67's value in gastric cancer (GC), we reevaluated the expression of Ki67 in whole tissue sections (WTS) and tissue microarrays (TMAs) of GC testing the following hypotheses: does Ki67 show intratumoral heterogeneity; are TMAs representative in the determination of the Ki67 proliferation index (PI); is the Ki67 PI subject to an intralaboratory variability; and is the Ki67 PI related to clinico-pathological patient characteristics and/or prognostically relevant in GC.
Methods: Corresponding WTS and TMAs samples from 315 GCs were stained immunohistochemically. The Ki67 PI evaluated on WTS was correlated with the Ki67 PI evaluated on TMAs, sample age, clinico-pathological characteristics, and patient survival.
Results: The overall amount of Ki67-positive tumor cells did not depend on sample age. Three distinct, partly heterogeneous Ki67 expression patterns were observed. The mean Ki67 PI evaluated on TMAs differed on average minus 16.9% from the Ki67 PI evaluated on WTS. Ki67 in WTS correlated significantly with the Laurén phenotype and tumor grade, but not with patient survival.
Conclusion: TMAs carry the risk of a systematic underestimation of the Ki67 PI. Ki67 has no prognostic value in GC but might be a potential indicator of intratumoral heterogeneity.
J. Surg. Oncol. 2016;113:46–54. © 2015 The Authors. *Journal of Surgical Oncology* published by Wiley Periodicals, Inc.

KEY WORDS: **proliferation; prognostic biomarker; tumor heterogeneity**

INTRODUCTION

Prognostic biomarkers are the most important instrument for tailoring oncologic treatment and predicting cancer patients prognosis. In gastric cancer (GC), the tumor- (T), node- (N), metastasis- (M) classification of the union internationale contre le cancer (UICC) is the single currently used prognostic biomarker on a routine basis, but recent studies have shown that the 7th edition of the UICC-stage grouping is not able to sufficiently discriminate patient's survival [1–3]. Additional reliable prognostic biomarkers, for example, immunohistochemistry or RNA/DNA-based tests, are urgently needed but not established yet on a routine basis [4].

The quest for a significant prognostic biomarker in GC has led to the examination of Ki67 over 20 years ago [5–7]. Ki67 is a nuclear located protein that is closely linked to cell proliferation. It is present in all active phases of the cell cycle, but absent from resting cells, thus, indicating the proliferating cell fraction [8]. Ki67 is an established prognostic biomarker in several tumor entities, for example, breast cancer, lymphoma, and neuroendocrine neoplasia [9–11]. Despite several previous investigations, the prognostic value of Ki67 proliferation index (PI) in GC remains contradictious. Different analytical methods to assess the Ki67 PI were used, and former investigations mainly relied on tissue microarrays (TMAs) without ever evaluating whether TMAs are suitable in general for the determination of the Ki67 PI in GC.

In order to fill this gap of information, we systemically investigated the prognostic value of the Ki67 PI in a large and thoroughly characterized cohort of GC, testing the following hypothesis: (i) Does Ki67 show any distinct intratumoral expression pattern in whole tissue sections (WTS); (ii) is the detection of Ki67 PI influenced by sample age, (iii) are TMAs applicable in the determination of the Ki67 PI in GC; and (iv) is the Ki67 PI related to clinico-pathological patient characteristics and/or prognostically relevant in GC.

MATERIALS AND METHODS

Study Population

From the archive of the Institute of Pathology, University Hospital Kiel, we sought caucasian patients who had undergone either total or partial gastrectomy for adenocarcinoma of the stomach or oesophago-gastric junction between 1997 and 2009. The following patient characteristics were retrieved: type of surgery, age at diagnosis, gender, tumor size, tumor localization, tumor type, tumor grade, depth of invasion, residual tumor status, number of lymph nodes resected, and number of lymph nodes with metastases. Patients were included if: (i) an adenocarcinoma of the stomach or oesophago-gastric junction was histologically confirmed; (ii) the overall tumor mass was large enough to get five TMA punches; and (iii) the Ki67 PI could be assigned on both WTS *and* corresponding TMAs obtained from the same paraffin blocks. Exclusion criteria were defined as: (i) histology identified a tumor type other than adenocarcinoma; and (ii) patients had undergone a perioperative chemo- or radiotherapy. Each resected specimen had

Grant sponsor: German Research Foundation; Grant number: Ro 1173/12.
Conflict of interest: CB and HMB declare that they have no conflict of interest.

*Correspondence to: Christoph Röcken, PhD, Department of Pathology, Christian-Albrechts-University, Arnold-Heller-Str. 3, Haus 14, D-24105 Kiel, Germany. Fax: +49(0)431-597-3462.
E-mail: christoph.roecken@uksh.de

Received 13 August 2015; Accepted 11 November 2015

DOI 10.1002/jso.24104

Published online 28 December 2015 in Wiley Online Library (wileyonlinelibrary.com).

undergone gross sectioning and histological examination by trained and board certified surgical pathologists. Date of patient death was obtained from the *Epidemiological Cancer Registry* of the state of Schleswig-Holstein, Germany. Follow-up data of those patients who were still alive were retrieved from hospital records and general practitioners. Ethical approval was obtained from the local ethical review board (D 453/10). All patient data were pseudonymized prior to study inclusion.

Histology

Tissue specimens were fixed in formalin and embedded in paraffin (FFPE). Deparaffinized sections were stained with hematoxylin and eosin. Histological re-examination of primary tissue sections was carried out for all cases to assure if inclusion criteria were confirmed. Tumors were classified according to the Laurén classification [12] and re-examined by two surgical pathologists. pTNM-stage of all study patients was determined according to the 7th edition of the UICC guidelines [13].

Tissue Microarray Construction

FFPE tissue samples were used to generate TMAs as described previously [14]. Briefly, five morphologically representative regions of the paraffin "donor" blocks (tumor) were chosen, and tissue cylinders of 1.5 mm diameter were punched from these areas. Afterwards, the tissue cylinders were inserted into a new "recipient" paraffin block using a custom-built instrument (Beecher Instruments, Silver Spring, MD). The new recipient paraffin blocks were warmed in a 60°C heating cabinet for 7 min to create a sufficient bond between the tumor tissue and the recipient block paraffin. 2.5 μm thick serial sections were obtained from the new recipient paraffin blocks, dried in a 60°C heating cabinet for 6 hr and stored in polystyrene slide storage boxes at 8°C until use.

Immunohistochemistry, Assessment of Microsatellite Instability and Detection of *Helicobacter pylori*- and Epstein-Barr-Virus-Infection

Immunohistochemical stainings were carried out with a Bondmax (Leica Biosystems, Wetzlar, Germany) automated slide staining system, using the Polymer Refine Detection Kit (Menarini Diagnostics, Berlin, Germany) and a monoclonal rabbit antibody, directed against Ki67 (clone SP6, Abcam, Cambridge, United Kingdom). Pretreatment was done with ER2 for 20 min. The antibody was diluted in antibody diluent (Zytomed Systems, Berlin, Germany) and applied in a 1:300 dilution. Immunohistochemical stainings of Her2/neu, the assessment of microsatellite instability using immunohistochemistry and molecular biology, and the evaluation of *Helicobacter pylori*- and Epstein-Barr-virus-infection was carried out as previously described [4,15].

Evaluation of Immunostaining

The nuclear expression of Ki67 was manually counted within 500 tumor cells for each tumor in the area of the highest density of Ki67-positive nuclei ("hot spot"). All nuclear staining was considered as "positive," irrespective of its intensity. The Her2/neu status was assessed as previously described [15].

Study Design

In a first step, the Ki67 PI of each tumor was evaluated immunohistochemically on WTS and TMAs of corresponding tumors, and the level of agreement was assigned. The intralaboratory variability was studied by correlating the staining results of the WTS and TMAs with sample age. Additionally, the staining results of the WTS were correlated with clinico-pathological characteristics, including gender, age, Laurén phenotype, mucin phenotype, localization, T- and N-category, lymphatic invasion (L-category), venous invasion (V-category), UICC-stage (7th edition), lymph node ratio (LNR), tumor grade, residual tumor status, *Helicobacter pylori*-, Epstein-Barr-virus (EBV)-, microsatellite instability (MSI)-, and Her2/neu-status, as well as survival data.

Statistical Analysis

Statistical analyses were done using SPSS 20.0 (IBM Corporation, New York). Ki67 raw values were correlated with clinico-pathological patient characteristics and survival data. For comparison purposes, the Ki67 PI evaluated by WTS was dichotomized at the median; patients below median were classified as "Ki67 low," patients with a Ki67 PI $\geq$ median were classified as "Ki67 high." The Ki67 PI was split into quartiles additionally. Moreover, a subgroup analysis of intestinal- and diffuse-type GCs according to the Laurén classification was done. The correlation between the Ki67 PI of WTS and TMAs was calculated by using Spearman's rho and Cohen's kappa. A kappa value of 0.20 was considered to be poor, of 0.21–0.40 to be fair, of 0.41–0.60 to be moderate, of 0.61–0.80 to be good, and of 0.81–1.00 to be very good. Median overall survival and tumor specific survival was determined using the Kaplan–Meier method, and the log-rank test was used to determine significance. We further applied a Cox regression to calculate the influence of raw Ki67 PI on survival. The significance of correlation between clinico-pathological parameters and the Ki67 PI groups evaluated by WTS was tested using Fisher's exact test. For parameters of ordinal scale (T-category, N-category, UICC stage) we applied Kendall's tau test instead. For the comparison between raw Ki67 PI values and clinico-pathological characteristics we calculated median values and 25%- and 75%-percentiles, respectively. Significance of differences between median values was assessed using the Kruskal-Wallis test. P values were derived from two-tailed tests. A $P \leq 0.05$ was considered statistically significant. No adjustments were made. To account for the effects of multiple testing, we applied the explorative Simes (Benjamini-Hochberg) procedure [16].

RESULTS

Three hundred and fifteen patients fulfilled all study criteria. The clinico-pathological patient characteristics are summarized in Table I. According to Laurén, an intestinal phenotype was found in 170 (54.0%), a diffuse type in 82 (26.0%), a mixed type in 23 (7.3%), and an unclassifiable type in 40 (12.7%) patients. Overall survival data was available in 304 of 315 cases (96.5%), tumor specific survival data in 281 of 315 cases (89.2%). Mean follow-up period was 17.6 months (range 0.07–142.7 months).

Ki67 Expression Patterns in WTS

In WTS, three different Ki67 expression patterns could be observed: 132 cases (41.9%) showed a homogeneous or "diffuse" distribution of Ki67-positive tumor cells; 66 cases (20.9%) showed a heterogeneous Ki67-expression pattern with more abundant Ki67-positive tumor cells at the tumor surface and/or at the invasion front; and 117 cases (37.1%) showed a heterogeneous, "clonal" distribution of Ki67-positive tumor cells (Fig. 1). According to Laurén, a "homogeneous/diffuse" expression pattern was observed significantly more often in the unclassified phenotype, a "heterogeneous/superficial ± invasion front" distribution significantly more often in the diffuse and mixed phenotype, and a "heterogeneous/clonal" expression pattern significantly more often in the intestinal and mixed phenotype ($P < 0.001$; Table II).

TABLE I. Clinico-Pathological Patient Characteristics and Correlation With the Ki67 Proliferation Index

				Raw Ki67 PI			Dichotomized Ki67 PI			Ki67 PI split into Quartiles				
Characteristic	valid [n]		Total [n (%)]	Median	25th–75th Percentile	*P*-Value	Ki67 low [n (%)]	Ki67 high [n (%)]	*P*-value	Ki67 first quartile [n (%)]	Ki67 second quartile [n (%)]	Ki67 third quartile [n (%)]	Ki67 fourth quartile [n (%)]	*P*-value
Gender	315	Female	113 (35.9)	67	(41–82)	0.141[a]	60 (53.1)	53 (46.9)	0.412[c]	35 (31.0)	25 (22.1)	30 (26.5)	23 (20.4)	0.092[d]
		Male	202 (64.1)	69	(51–84)		97 (48.0)	105 (52.0)		43 (21.3)	54 (26.7)	49 (24.3)	56 (27.7)	
Age	315	≤68 years	142 (46.6)	68	(45–83)	0.544[a]	72 (50.7)	70 (49.3)	1.000[c]	40 (28.2)	32 (22.5)	37 (26.1)	33 (23.2)	0.454[d]
		>68 years	163 (53.4)	69	(51–84)		82 (50.3)	81 (49.7)		37 (22.7)	45 (27.6)	37 (22.7)	44 (27.0)	
Laurén phenotype	315	Intestinal	170 (54.0)	72	(55–85)	**<0.001**[a]	70 (41.2)	100 (58.8)	**<0.001**[c]	28 (16.5)	42 (24.7)	50 (29.4)	50 (29.4)	**0.004**[c]
		Diffuse	82 (26.0)	52	(35–75)		57 (69.5)	25 (30.5)		35 (42.7)	22 (26.8)	13 (15.9)	12 (14.6)	
		Mixed	23 (7.3)	72	(52–85)		10 (43.5)	13 (56.5)		5 (21.7)	5 (21.7)	6 (26.1)	7 (30.4)	
		Unclassified	40 (12.7)	69	(43–84)		20 (50.0)	20 (50.0)		10 (25.0)	10 (25.0)	10 (25.0)	10 (25.0)	
Mucin phenotype	279	Intestinal	82 (29.4)	74	(52–84)	0.241[a]	33 (40.2)	49 (59.8)	0.221[c]	19 (23.2)	14 (17.1)	26 (31.7)	23 (28.0)	0.114[c]
		Gastric	35 (12.5)	63	(40–80)		18 (51.4)	17 (48.6)		12 (34.3)	6 (17.1)	11 (31.4)	6 (17.1)	
		Mixed	117 (41.9)	68	(53–84)		61 (52.1)	56 (47.9)		25 (21.4)	36 (30.8)	24 (20.5)	32 (27.4)	
		Unclassified	45 (16.1)	59	(46–78)		26 (57.8)	19 (42.2)		12 (26.7)	14 (31.1)	13 (28.9)	6 (13.3)	
Localization	315	Proximal	108 (34.3)	76	(58–86)	0.003[a]	43 (39.8)	65 (60.2)	0.013[d]	21 (19.4)	22 (20.4)	30 (27.8)	35 (32.4)	0.009[d]
		Distal	207 (65.7)	65	(45–80)		114 (55.1)	93 (44.9)		57 (27.5)	57 (27.5)	49 (23.7)	44 (21.3)	
T-category	315	T1a/b	25 (7.9)	78	(57–88)	0.506[a]	9 (36.0)	16 (64.0)	0.251[d]	5 (20.0)	4 (16.0)	6 (24.0)	10 (40.0)	0.316[d]
		T2	44 (14.0)	72	(54–83)		19 (43.2)	25 (56.8)		9 (20.5)	10 (22.7)	17 (38.6)	8 (18.2)	
		T3	127 (40.3)	66	(47–83)		68 (53.5)	59 (46.5)		32 (25.2)	36 (28.3)	28 (22.0)	31 (24.4)	
		T4a/b	119 (37.8)	68	(45–84)		61 (51.3)	58 (48.7)		32 (26.9)	29 (24.4)	28 (23.5)	30 (25.2)	
N-category	314	N0	78 (24.8)	72	(54–86)	0.128[a]	30 (38.5)	48 (61.5)	0.017[d]	19 (24.4)	11 (14.1)	24 (30.8)	24 (30.8)	0.048[d]
		N1	47 (15.0)	73	(47–85)		21 (44.7)	26 (55.3)		12 (25.5)	9 (19.1)	10 (21.3)	16 (34.0)	
		N2	57 (18.2)	64	(48–84)		32 (56.1)	25 (43.9)		14 (24.6)	18 (31.6)	10 (17.5)	15 (26.3)	
		N3/a/b	132 (42.0)	65	(48–80)		73 (55.3)	59 (44.7)		33 (25.0)	40 (30.3)	35 (26.5)	24 (18.2)	
L-category	303	L0	141 (46.5)	68	(47–83)	0.744[a]	73 (51.8)	68 (48.2)	0.491[c]	37 (26.2)	36 (25.5)	36 (25.5)	32 (22.7)	0.367[d]
		L1	162 (53.5)	69	(49–84)		77 (47.5)	85 (52.5)		38 (23.5)	39 (24.1)	41 (25.3)	44 (27.2)	
V-category	301	V0	263 (87.4)	69	(47–83)	0.778[a]	131 (49.8)	132 (50.2)	1.000[c]	67 (25.5)	64 (24.3)	67 (25.5)	65 (24.7)	0.805[d]
		V1	38 (12.6)	70	(54–84)		19 (50.0)	19 (50.0)		8 (21.1)	11 (28.9)	9 (23.7)	10 (26.3)	
UICC Stage (7th ed.)	310	IA	18 (5.8)	75	(40–87)	0.024[a]	7 (38.9)	11 (61.1)	0.009[d]	5 (27.8)	2 (11.1)	5 (27.8)	6 (33.3)	0.012[d]
		IB	21 (6.8)	74	(54–89)		8 (38.1)	13 (61.9)		4 (19.0)	4 (19.0)	7 (33.3)	6 (28.6)	
		IIA	42 (13.5)	73	(51–86)		16 (38.1)	26 (61.9)		10 (23.8)	6 (14.3)	12 (28.6)	14 (33.3)	
		IIB	31 (10.0)	70	(47–84)		16 (51.6)	15 (48.4)		10 (32.3)	6 (19.4)	8 (25.8)	7 (22.6)	
		IIIA	44 (14.2)	66	(51–84)		23 (52.3)	21 (47.7)		8 (18.2)	15 (34.1)	10 (22.7)	11 (25.0)	
		IIIB	52 (16.8)	69	(55–86)		25 (48.1)	27 (51.9)		11 (21.2)	14 (26.9)	10 (19.2)	17 (32.7)	
		IIIC	43 (13.9)	69	(53–81)		21 (48.8)	22 (51.2)		8 (18.6)	13 (30.2)	13 (30.2)	9 (20.9)	
		IV	59 (19.0)	56	(40–73)		38 (64.4)	21 (35.6)		22 (37.3)	16 (27.1)	14 (23.7)	7 (11.9)	
LNR	306	≤0.229	153 (50.0)	73	(54–86)	0.005[a]	64 (41.8)	89 (58.2)	0.016[c]	33 (21.6)	31 (20.3)	39 (25.5)	50 (32.7)	0.008[d]
		>0.229	153 (50.0)	65	(45–78)		86 (56.2)	67 (43.8)		40 (26.1)	46 (30.1)	39 (25.5)	28 (18.3)	
Tumor grade	307	G1/G2	70 (22.8)	74	(53–90)	**0.005**[a]	26 (37.1)	44 (62.9)	0.015[c]	14 (20.0)	12 (17.1)	20 (28.6)	24 (34.3)	0.020[d]
		G3/G4	237 (77.2)	65	(46–81)		128 (54.0)	109 (46.0)		62 (26.2)	66 (27.8)	58 (24.5)	51 (21.5)	
Residual tumor	302	R0	263 (87.1)	69	(50–83)	0.528[a]	128 (48.7)	135 (51.3)	0.864[c]	62 (23.6)	66 (25.1)	70 (26.6)	65 (24.7)	0.994[d]
		R1/R2	39 (12.9)	63	(38–86)		20 (51.3)	19 (48.7)		11 (28.2)	9 (23.1)	7 (17.9)	12 (30.8)	
Helicobacter pylori status	266	Negative	221 (83.1)	68	(49–83)	0.724[a]	113 (51.1)	108 (48.9)	0.626[c]	55 (24.9)	58 (26.2)	55 (24.9)	53 (24.0)	0.524[d]
		Positive	45 (16.9)	71	(52–86)		21 (46.7)	24 (53.3)		10 (22.2)	11 (24.4)	11 (24.4)	13 (28.9)	
EBV status	308	Negative	296 (96.1)	68	(49–84)	0.569[a]	150 (50.7)	146 (49.3)	0.138[c]	73 (24.7)	77 (26.0)	71 (24.0)	75 (25.3)	0.444[d]
		Positive	12 (3.9)	74	(56–80)		3 (25.0)	9 (75.0)		1 (8.3)	2 (16.7)	8 (66.7)	1 (8.3)	
MSI	303	MSS	278 (91.7)	68	(50–83)	0.346[a]	140 (50.4)	138 (49.6)	0.211[c]	66 (23.7)	74 (26.6)	69 (24.8)	69 (24.8)	0.360[d]
		MSI-H	25 (8.3)	72	(63–84)		9 (36.0)	16 (64.0)		5 (20.0)	4 (16.0)	9 (36.0)	7 (28.0)	
Her2/neu status	315	Negative	289 (91.7)	67	(47–83)	0.011[a]	149 (51.6)	140 (48.4)	0.064[c]	74 (25.6)	75 (26.0)	70 (24.2)	70 (24.2)	0.076[d]
		Positive	26 (8.3)	81	(66–91)		8 (30.8)	18 (69.2)		4 (15.4)	4 (15.4)	9 (34.6)	9 (34.6)	
Overall survival [months]	304	Events (Dead)	243 (79.9)			0.274[b]	125 (81.2)	118 (78.7)	0.564[e]	63 (81.8)	62 (80.5)	60 (81.1)	58 (76.3)	0.758[e]
		Alive	61 (20.1)				29 (18.8)	32 (21.3)		14 (18.2)	15 (19.5)	14 (18.9)	18 (23.7)	
		Median survival [Months]					14.0 ± 1.6	14.1 ± 1.7		15.0 ± 2.0	13.2 ± 2.0	13.8 ± 1.8	14.7 ± 2.5	
		95%CI					10.9–17.1	10.7–17.4		11.1–18.8	9.3–17.1	10.4–17.3	9.7–19.6	

TABLE I. (*Continued*)

Characteristic	valid [n]		Total [n (%)]	Raw Ki67 PI			Dichotomized Ki67 PI			Ki67 PI split into Quartiles				
				Median	25th–75th Percentile	*P*-Value	Ki67 low [n (%)]	Ki67 high [n (%)]	*P*-value	Ki67 first quartile [n (%)]	Ki67 second quartile [n (%)]	Ki67 third quartile [n (%)]	Ki67 fourth quartile [n (%)]	*P*-value
Tumor specific survival [months]	281	Events (Dead)	197 (70.1)			0.196[b]	102 (72.3)	95 (67.9)	0.370[e]	49 (71.0)	53 (73.6)	47 (69.1)	48 (66.7)	0.731[e]
		Alive/ DOOD	84 (29.9)				39 (27.7)	45 (32.1)		20 (29.0)	19 (26.4)	21 (30.9)	24 (33.3)	
		Median survival [Months]					16.0 ± 1.8	15.6 ± 2.3		16.7 ± 1.8	13.6 ± 2.7	15.4 ± 2.8	15.6 ± 2.1	
		95%CI					12.5–19.5	11.2–20.1		13.1–20.2	8.2–18.9	9.9–20.8	8.5–24.7	

P-values printed in bold denote significant correlation after Simes' multiple testing procedure.
UICC, union internationale contre le cancer; LNR, lymph node ratio; EBV, Epstein-Barr-virus; MSI, microsatellite instability; DOOD, died of other disease.
[a]Kruskal-Wallis test.
[b]Cox regression. HR(Overall survival)=0.997 [0.991-1.002]; HR(Tumor specific survival)=0.996 [0.990–1.002].
[c]Fisher's exact test.
[d]Kendall's tau test.
[e]log-rank test.

Fig. 1. Ki67 shows a heterogeneous expression pattern in the majority of gastric cancers. (**a–c**) shows a case with an increased Ki67 expression at the tumor surface (**b**), whereas deep-lying areas show a low Ki67 PI (**c**). (**d**) illustrates a gastric carcinoma with a heterogeneous, "clonal" Ki67 expression pattern, characterized by tumor areas with a very high Ki67 PI (right) that are located in close proximity to tumor portions with a very low Ki67 PI (left). (**e**) and (**f**) are taken from the WTS (**e**) respectively the TMA (**f**) of the same tumor that shows a homogeneous, "diffuse" distribution of Ki67 positive tumor cells and, accordingly, a good level of agreement between the Ki67 PI evaluated on WTS and TMAs. Ki67 immunostaining, original magnification 50-fold (**a**), 250-fold (**b–f**).

Ki67 Pi in WTS and TMAs

The Ki67 PI assessed by WTS ranged from 2% to 99% (mean 64.3%, median 69.0%, standard deviation 22.4%). Divided by the median, 157 cases (49.8%) were classified as "Ki67 low," and 158 cases (50.2%) were classified as "Ki67 high." The first quartile contained 78 cases

TABLE II. Correlation Between the Laurén Phenotype and Ki67 Expression Patterns

Characteristic	valid [n]		Total [n (%)]	Ki67 expression pattern: Homogeneous/"diffuse" [n (%)]	Ki67 expression pattern: Heterogeneous/superficial ± invasion front [n (%)]	Ki67 expression pattern: Heterogeneous/"clonal" [n (%)]	*P*-value
Laurén phenotype	315	Intestinal	170 (54.0)	62 (36.5)	30 (17.6)	78 (45.9)	<0.001[a]
		Diffuse	82 (26.0)	40 (48.8)	21 (25.6)	21 (25.6)	
		Mixed	23 (7.3)	3 (13.0)	8 (34.8)	12 (52.2)	
		Unclassified	40 (12.7)	27 (67.5)	7 (17.5)	6 (15.0)	

[a]Fisher's exact test.

(24.7%), the second, third, and fourth quartile contained each 79 cases (each 25.1%).

The Ki67 PI assessed by TMAs ranged from 1% to 99% (mean 47.4%, median 44.5%, standard deviation 26.6%). Divided by the Ki67 PI median evaluated on WTS, 231 cases (73.3%) were classified as "Ki67 low," and 84 cases (26.7%) were classified as "Ki67 high."

Spearman's rho implied a positive agreement between the WTS- and TMA-Ki67 PI ($r_s = 0.615$; Fig. 2). Split by the median, both groups

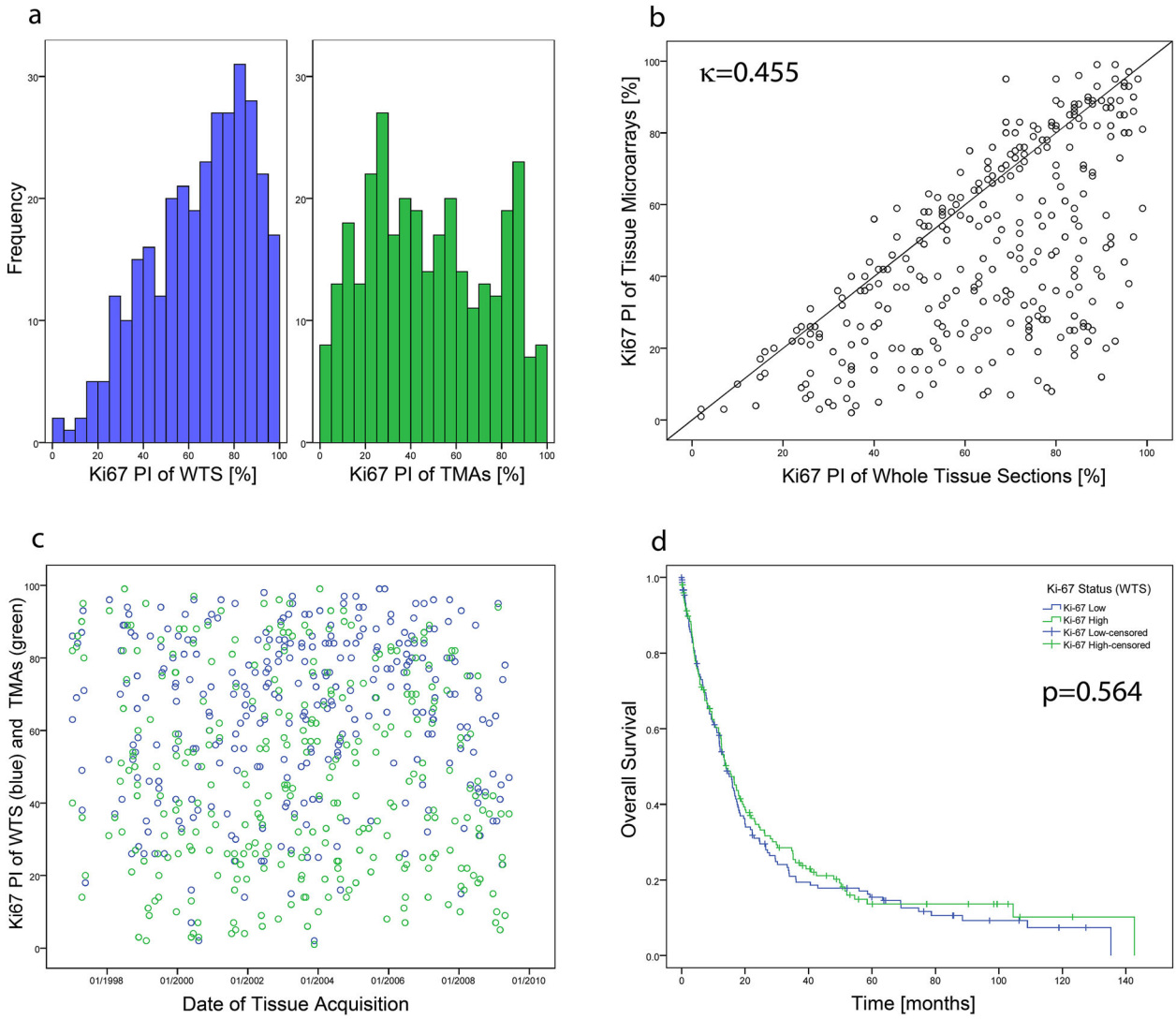

Fig. 2. Agreement between the Ki67 proliferation index evaluated on WTS and on TMAs. Correlation between the Ki67 proliferation index (Ki67 PI) evaluated on WTS and on TMAs for the entire cohort (**a**, **b**). Both groups showed a moderate measure of agreement ($\kappa = 0.455$; **b**). The Ki67 PI in our cohort did not correlate with the sample age of WTS (blue) or TMAs (green) (**c**). There was no significant correlation between the Ki67 PI and overall patient survival ($P = 0.564$; **d**).

showed a moderate measure of agreement [$\kappa = 0.455$; asymptomatic standard error (ASE) 0.044].

Regarding the Laurén phenotype, the level of agreement was moderate for intestinal ($\kappa = 0.457$; ASE 0.057), diffuse ($\kappa = 0.412$; ASE 0.106) and unclassified GCs ($\kappa = 0.450$; ASE 0.137), and fair for mixed type GCs ($\kappa = 0.279$; ASE 0.130).

Evaluating the correlation between the Ki67 expression pattern and the level of agreement of the Ki67 PI in WTS and TMAs, carcinomas with a "homogeneous/diffuse" Ki67 expression pattern showed a good level of agreement ($\kappa = 0.673$; ASE 0.064), whereas tumors with a heterogeneous Ki67 expression pattern showed a fair level of agreement (Ki67-positive cells superficial and/or at invasion front: $\kappa = 0.367$, ASE 0.104; "clonal" expression: $\kappa = 0.296$, ASE 0.061; Supplemental File S1).

In view of the moderate agreement between TMA and WTS, we carried out all subsequent analyses using WTS only.

Evaluation of Intralaboratory Variability

Next we assessed the putative influence of sample age on the amount of Ki67-positive tumor cells in WTS and TMAs. Using Spearman's rho, no significant agreement between the sample age and the Ki67 PI in WTS ($r_s = 0.058$, $P = 0.302$) or TMAs ($r_s = 0.202$, $P = 0.127$) was found (Fig. 2).

Clinico-Pathological Correlation

Ki67 raw values evaluated on WTS were correlated with clinico-pathological patient characteristics. Moreover, the Ki67 PI was split by the median, and all cases were divided accordingly into "Ki67 low" ($<$median) and "Ki67 high" ($\geq$median); additionally, the Ki67 PI was divided into quartiles. Ki67 correlated significantly with various clinico-pathological patient characteristics (complete data is shown in Table I and in Supplemental File S2). After Simes' multiple testing procedure, the Laurén-phenotype and the tumor grade remained the only clinico-pathological parameter which correlated significantly with the Ki67 PI: The Ki67 PI was significantly higher in intestinal- and mixed-type GCs compared to unclassified and diffuse-type GCs ($P < 0.001$ for raw and dichotomized Ki67 values). Moreover, the median Ki67 PI (raw value) was significantly higher in G1/G2 differentiated GCs, compared to G3/G4 GCs (74% vs. 65%, $P = 0.005$). There was no other significant correlation between the Ki67 PI and any other clinico-pathological patient characteristic after Simes' multiple testing procedure.

Prognostic Significance

Patient prognosis significantly depended on the Laurén phenotype, tumor grade, T-, N-, L-, and V-category, LNR, as well as UICC-stage, residual tumor status, and MSI status (Supplemental File S3). There was no significant correlation between the Ki67 status and overall ($P = 0.564$) or tumor specific patient survival ($P = 0.370$; Fig. 2).

Subgroup Analysis

The subgroup analysis of intestinal- and diffuse-type GCs revealed that patients with an intestinal-type GC and 0–2 lymph node metastases (N0-/N1-category) respectively a LNR <0.229 had a significant higher median Ki67 PI than patients with $\geq$3 lymph node metastases (N2-/N3-category; $P = 0.029$) respectively a LNR ≥ 0.229 ($P = 0.009$). Additionally, a higher median Ki67 was observed significantly more often in UICC-stage I and II intestinal-type GCs compared to UICC-stage III and IV intestinal-type GCs ($P = 0.007$). Moreover, patients with a diffuse-type GC and a Her2/neu-overexpression/amplification had a significant higher median Ki67 PI compared to Her2/neu-negative diffuse-type GCs (82.0% vs. 52.6%; $P = 0.036$). However, in the subgroup analysis of intestinal- and diffuse-type GCs, the Ki67 PI did not correlate significantly with any of the tested clinico-pathological parameters nor with patient survival after Simes' multiple testing procedure (data not shown).

DISCUSSION

Gastric cancer is a prevalent, aggressive, and heterogeneous disease with a poor prognosis [17]; its distinct genetic complexity has been recently shown in an integrative genomic analysis including whole-genome sequencing. A molecular classification of GC was proposed, which categorizes four subtypes: Epstein-Barr-virus (EBV)-positive, microsatellite instable (MSI), chromosomal instable, and genomically stable GCs [18,19]. Additionally, evidence has accumulated indicating that patient prognosis and treatment-response depends not only on the UICC tumor stage, but also on the expression and tumor specific alterations of intracellular signaling pathways [20]. However, with regard to patient management, risk stratification is a major issue and various assays have been exploited in many different types of cancer, including gene expression profiling and counting mitoses. With regard to GC, except for TNM-classification and the assessment of LNR, no other tissue based marker has stood the test of time or reached clinical practice.

One of the oldest, most widely and sometimes most uncritically used, tissue based marker of prognosis is Ki67. Firstly, reported in 1983 and named after its city of origin (Kiel, Germany) and the number of the original clone in the 96-well plate, the Ki67 antibody recognizes a nuclear antigen present in proliferating cells, which is absent in resting cells [21]. Since then and despite great efforts, Ki67 could be established only in a minority of cancer entities because it often led to inconsistent and disappointing results. This also applies to GC. Ki67 has been tested as a prognostic biomarker in GC several times during the last decades, but the results remained contradictory: Several authors described a correlation between a *high* Ki67 PI and a poor prognosis, whereas Lee et al. [22] associated a *low* Ki67 PI with a poor patient's outcome; still other authors declared that Ki67 is of no prognostic value at all. These discrepancies might be related to methodological issues: The number of patients studied in 13 former studies ranged from 43 to 418. Apart from one single study, all other larger series relied on the analysis of TMAs, without ever testing whether this is a reliable approach to assess the Ki67 PI in GC at all. The number of cells counted in a single tumor ranged by a factor of ten from 100 to 1,000 or was sometimes not even indicated. Moreover, the cut-off value to determine "Ki67 high" seemed to be chosen arbitrarily in most studies (Table III) [5,22–33].

Thus, the vast majority of former studies, which have investigated the significance of Ki67 PI in GC might have been biased by methodological issues. In the present study, we intended to circumvent this problem and explored for the first time systematically the value of TMAs for the assessment of KI67 PI in GC by comparing staining with WTS obtained from the same paraffin blocks.

With regard to Ki67 scoring standardization, it is generally difficult to determine a single useful cut-off point: Firstly, Ki67 displays a continuous distribution; secondly, Ki67 is highly capable of being influenced by pre-analytical and analytical variables. In breast cancer, where Ki67 is routinely used as a part of a multi-parameter prognostic biomarker, substantial variability in Ki67 scoring is known to be present among some of the world's most experienced laboratories. As a consequence, Ki67 values and cut-offs cannot be transferred between different laboratories without standardizing scoring methodology [34,35]. It is recommended that Ki67 scores should be interpreted in the light of local laboratory values, instead of suggesting a universally valid cut-off value [36]. Consequently, we abstained from applying defined cut-off values as previously done in many other studies, and chose an evaluation of the raw Ki67 values and additionally a separation by the median and into quartiles instead, as this is the most reliable way to consider interlaboratory variability.

TABLE III. Literature Review

Reference	Publication date	n	Sample material	Cells counted	Mean %	Median %	Cut-off	Correlation with clinco-pathological parameters	Correlation with survival
Müller et al.	1996	418	WTS	1000	51.3	53.3	Median	None	None
Ramires et al.	1997	43	WTS	1000	32.7	n.s.	n.s.	None	n.e.
Manzoni et al.	1998	56	WTS	1000	n.s.	n.s.	<10%, 10–40%, >40%	Higher Ki67 PI in patients >68 years	High Ki67 PI correlates with poor prognosisin patients >68 years
Oshima et al.	2005	70	WTS	1000	n.s.	n.s.	40%	None	None
Czyzewska et al.	2009	100	WTS	10 HPF	n.s.	n.s.	50%	High Ki67 PI in pT3/4 respectively pN2 tumors	None
Tsamandas et al.	2009	110	WTS	500	25.3	30.0	Median	High Ki67 PI in G3/G4 tumors	High Ki67 PI correlates with poor prognosis
Lee et al.	2010	245	TMA	300	23.1	15.7	0–10%, 11–20%, 21–30%. 31–40%,41–50%, >50%	High Ki67 in G1/G2, pT1/2, pN0 tumors	High Ki67 correlates with good prognosis
Lazar et al.	2010	67	WTS	500	46.2	n.s.	45%	High Ki67 in patients >60 years, cardia tumors, papillary adenocarcinoma, G3 tumors	None
He et al.	2012	166	TMA	1000	n.s.	n.s.	0–4%, 5–25%, 26–50%, >50%	n.e.	High Ki67 PI correlates with poor prognosis
Kroepil et al.	2013	163	TMA	n.s.	n.s.	7	Median	None	n.e.
Xiao et al.	2013	413	TMA	100	n.s.	n.s.	0%, 1–50%, 50–74%, 75–100%	High Ki67 in UICC stage I tumors	None
Wu et al.	2014	101	WTS	n.s.	n.s.	n.s.	0–10%, 11–50%, 51–75%, >75%	High Ki67 in poorly differentiate tumors, tumors with lymph node metastasis, UICC stage III-/IV-tumors, Her2/neu-positive tumors	n.e.
Saricanbaz et al.	2014	50	WTS	n.s.	n.s.	n.s.	0–5%, 6–30%, 30–60%, 61–100%	None	None

WTS, whole tissue sections; TMA, tissue microarray; n.s., not specified; n.e., not evaluated.

Ki67 Is a Potential Indicator of Intratumoral Heterogeneity

Intratumoral genetic and phenotypic heterogeneity is another major risk for sampling errors, as smaller tissue samples, for example, biopsies, bear the risk of being not representative for the entire tumor mass. Although, it is well known in GC, for example, in the evaluation of the Her2/neu status, it is widely neglected in most biomarker studies [15,37]. Heterogeneity may not solely be present among cases of a particular tumor type, but also within an individual cancer with subclones that might response variably to anti-cancer drugs. Thus, biomarkers with the ability to indicate intratumoral heterogeneity are urgently needed. Interestingly, Ramires et al. already described distinct intratumoral Ki67 expression patterns in GC that correlated significantly with the Laurén phenotype. They observed that subgroups of GCs showed an increased Ki67 expression within the tumor surface compared to deeper tumor areas, and vice versa. Moreover, a "focal clustering" was also described in a subset of GCs [23]. In addition to these previous findings, we noticed that three distinct Ki67 expression patterns are present in GC, and that two of them showed a heterogeneous Ki67-expression pattern that might indicate intratumoral heterogeneity.

TMAs Are Unsuitable for the Ki67 PI Evaluation in GC

In our patient series, only the minority of GCs showed an even, homogeneous intratumoral distribution of Ki67-positive tumor cells. We made similar observations with regard to heterogeneous expression of HER2 and MET in GC and came to the conclusion that TMAs, like any other tissue biopsy, carry the risk of a sampling error [15,38]. This risk is exemplified by: (i) the fact that the mean Ki67 PI evaluated on TMAs underestimated the Ki67 PI by 16.9%; and (ii) the moderate level of agreement between TMA and WTS, which raised to a good level of agreement if those tumors with a "homogeneous/diffuse" Ki67 expression pattern were evaluated separately. These findings lead to the conjecture that studies using TMAs only, carry the risk of a sampling error due to a non-representative evaluation and systematic underestimation of the Ki67 PI in GC.

The Ki67 PI Does Not Depend on Sample Age

Collecting sufficient patient numbers is time consuming, and storage of paraffin blocks and changes in pre-analytical procedures may lead to antigen degradation/modification influencing test results [39]. The tissue samples of our study were collected over a period of 12 years and had been stored between 6 and 18 years prior to study execution. However, in our series no correlation was found between sample age and neither quantity nor intensity of Ki67 immunostaining, neither for WTS nor for TMAs. Thus, sample age does not influence the test result of the KI67 PI.

The Ki67 PI Is Correlated Significantly With the Laurén Phenotype and Tumor Grade, But Has No Prognostic Value in GC

The conclusion that a high amount of proliferating tumor cells indicates a fast growing tumor with an increased potential to metastasize and a worse prognosis has been recently discussed controversially [40]. Contradictory to most previous studies, Lee et al. stated that highly proliferative tumors possess less invasive subclones and thus, have a less aggressive metastatic potential that

results in a better patient prognosis [22]. Nevertheless, Lee's findings are compromised by their use of TMAs and thereby may have missed intratumoral heterogeneity of GC. Neither previous finding, that is, positive or negative prediction of patient prognosis, could be confirmed in our cohort.

Interestingly, the Ki67 PI correlated with the Laurén phenotype and the tumor grade, and a significant correlation between the distinct expression patterns and the Laurén phenotype was also observed. The molecular classification of GC [18] also correlates to some extent with the Laurén phenotype: Genomically stable GCs were enriched in the diffuse type GC, which also showed most commonly a homogeneous Ki67 PI in our cohort. The chromosomal instable GCs were more commonly associated with an intestinal phenotype, which in our analysis most frequently harbored a heterogeneous/clonal Ki67 distribution. Chromosomal instable GCs often show amplification of receptor tyrosine kinases and our previous studies on the same cohort confirmed the occurrence of intratumoral heterogeneity with *HER2* or *MET*-amplified and unamplified tumor cell clones, respectively, within the same patient [15,38]. Thus, Ki67 PI in GC may be linked to the Laurén phenotype based on molecular biological differences (genomically stable vs. genomically instable) rather than indicators of patient prognosis. Indeed, Ki67 PI assessed on WTS did not correlate with patient survival at all in our series, neither for the entire cohort nor for the subgroup analysis of intestinal- and diffuse-type GCs.

While interpreting our findings, it should be kept in mind that the proliferation rate of a tumor is a temporal snap-shot and may provide no information about the history or future development of the individual cancer. In future studies, the presence of the three distinct Ki67 expression patterns needs to be correlated with other indicators of intratumoral heterogeneity, especially molecular characteristics.

Summing up, the present study is the first that considers (and eliminates) the risks of sampling errors when using TMAs (or biopsies) of GCs. The assessment of Ki67 in archival tissue samples is not influenced by intralaboratory pre-analytical variables (e.g., storage time) and is unsuitable as a prognostic biomarker for GC. However, Ki67 may unravel tumor heterogeneity, and future studies should pay special attention to a standardized evaluation and appropriate and representative tissue sampling.

AUTHORS' CONTRIBUTIONS

CB conceived and carried out experiments and analyzed data. CR conceived experiments. HMB analyzed data. All authors were involved in writing the paper and had final approval of the submitted manuscript.

ETHICS

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and later versions. Informed consent or substitute for it was obtained from all patients for being included in the study. Ethical approval was obtained from the local ethical review board (D 453/10).

ACKNOWLEDGMENTS

We wish to thank Ann-Kristin Sternberg and Deniza Hajzeri for their excellent technical assistance. CR is supported by grants of the German Research Foundation (Ro 1173/12).

REFERENCES

1. Röcken C, Behrens HM: Validating the prognostic and discriminating value of the TNM-classification for gastric cancer—A critical appraisal. Eur J Cancer 2015;51:577–586.
2. Patel MI, Rhoads KF, Ma Y, et al.: Seventh edition (2010) of the AJCC/UICC staging system for gastric adenocarcinoma: is there room for improvement? Ann Surg Oncol 2013;20:1631–1638.
3. Yoon HM, Ryu KW, Nam BH, et al.: Is the new seventh AJCC/UICC staging system appropriate for patients with gastric cancer? J Am Coll Surg 2012;214:88–96.
4. Warneke VS, Behrens HM, Haag J, et al.: Prognostic and putative predictive biomarkers of gastric cancer for personalized medicine. Diagn Mol Pathol 2013;22:127–137.
5. Muller W, Schneiders A, Meier S, et al.: Immunohistochemical study on the prognostic value of MIB-1 in gastric carcinoma. Br J Cancer 1996;74:759–765.
6. Porschen R, Loh B, Hengels KJ, et al.: Assessment of cell proliferation in colorectal carcinomas using the monoclonal antibody Ki-67. Correlation with pathohistologic criteria and influence of irradiation. Cancer 1989;64:2501–2505.
7. Yonemura Y, Ooyama S, Sugiyama K, et al.: Growth fractions in gastric carcinomas determined with monoclonal antibody Ki-67. Cancer 1990;65:1130–1134.
8. Scholzen T, Gerdes J: The Ki-67 protein: From the known and the unknown. J Cell Physiol 2000;182:311–322.
9. Stuart-Harris R, Caldas C, Pinder SE, et al.: Proliferation markers and survival in early breast cancer: A systematic review and meta-analysis of 85 studies in 32,825 patients. Breast 2008;17:323–334.
10. He X, Chen Z, Fu T, et al.: Ki-67 is a valuable prognostic predictor of lymphoma but its utility varies in lymphoma subtypes: Evidence from a systematic meta-analysis. BMC Cancer 2014;14:153.
11. van Velthuysen ML, Groen EJ, van der Noort V, et al.: Grading of neuroendocrine neoplasms: Mitoses and Ki-67 are both essential. Neuroendocrinology 2014;100:221–227.
12. Lauren P: The two histological main types of gastric carcinoma—Diffuse and so-called intestinal-type carcinoma—An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 1965;64:31–49.
13. Wittekind C, Oberschmid B: TNM classification of malignant tumors 2010. Pathologe 2010;31:333–338.
14. Packeisen J, Korsching E, Herbst H, et al.: Demystified... Tissue microarray technology. Mol Pathol 2003;56:198–204.
15. Warneke VS, Behrens HM, Boger C, et al.: Her2/neu testing in gastric cancer: Evaluating the risk of sampling errors. Ann Oncol 2013;24:725–733.
16. Benjamini Y: Discovering the false discovery rate. J R Stat Soc Series B Stat Methodol 2010;72:405–416.
17. Guggenheim DE, Shah MA: Gastric cancer epidemiology and risk factors. J Surg Oncol 2013;107:230–236.
18. The Cancer Genome Atlas Research Network: Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014;513:202–209.
19. Wang K, Yuen ST, Xu J, et al.: Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. Nat Genet 2014;46:573–582.
20. Lordick F, Schulze T: Molecular prognostic factors and new systemic therapies in gastric cancer. Onkologe 2008;14:389–395.
21. Gerdes J, Schwab U, Lemke H: Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. Int J Cancer 1983;31:13–20.
22. Lee HE, Kim MA, Lee BL: Low Ki-67 proliferation index is an indicator of poor prognosis in gastric cancer. J Surg Oncol 2010;102:201–206.
23. Ramires M, David L, Leitao D, et al.: Ki67 labelling index in gastric carcinomas. An immunohistochemical study using double staining for the evaluation of the proliferative activity of diffuse-type carcinomas. J Pathol 1997;182:62–67.
24. de Manzoni G, Verlato G, Tomezzoli A, et al.: Study on Ki-67 immunoreactivity as a prognostic indicator in patients with advanced gastric cancer. Jpn J Clin Oncol 1998;28:534–537.
25. Oshima CT, Iriya K, Forones NM: Ki-67 as a prognostic marker in colorectal cancer but not in gastric cancer. Neoplasma 2005;52:420–424.

26. Czyzewska J, Guzinska-Ustymowicz K, Pryczynicz A, et al.: Immunohistochemical evaluation of Ki-67, PCNA, and MCM2 proteins proliferation index (PI) in advanced gastric cancer. Folia Histochem Cytobiol 2009;47:289–296.
27. Tsamandas AC, Kardamakis D, Tsiamalos P, et al.: The potential role of Bcl-2 expression, apoptosis and cell proliferation (Ki-67 expression) in cases of gastric carcinoma and correlation with classic prognostic factors and patient outcome. Anticancer Res 2009;29:703–709.
28. Lazar D, Taban S, Sporea I, et al.: Ki-67 expression in gastric cancer. Results from a prospective study with long-term follow-up. Rom J Morphol Embryol 2010;51:655–661.
29. He WL, Li YH, Yang DJ, et al.: Combined evaluation of centromere protein H and Ki-67 as prognostic biomarker for patients with gastric carcinoma. Eur J Surg Oncol 2013;39: 141–149.
30. Kroepil F, Dulian A, Vallbohmer D, et al.: High EpCAM expression is linked to proliferation and lauren classification in gastric cancer. BMC Res Notes 2013;6:253.
31. Xiao LJ, Zhao S, Zhao EH, et al.: Clinicopathological and prognostic significance of Ki-67, caspase-3, and p53 expression in gastric carcinomas. Oncol Lett 2013;6:1277–1284.
32. Wu HW, Qin CY, Huang JL, et al.: Correlations of beta-catenin, Ki67, and Her-2/neu with gastric cancer. Asian Pac J Trop Med 2014;7:257–261.
33. Saricanbaz I, Karahacioglu E, Ekinci O, et al.: Prognostic significance of expression of CD133 and Ki-67 in gastric cancer. Asian Pac J Cancer Prev 2014;15:8215–8219.
34. Polley MY, Leung SC, McShane LM, et al.: An international Ki67 reproducibility study. J Natl Cancer Inst 2013;105:1897–1906.
35. Polle MY, Leung SC, Gao D, et al.: An international study to increase concordance in Ki67 scoring. Mod Pathol 2015;28: 778–786.
36. Coates AS, Winer EP, Goldhirsch A, et al.: Tailoring therapies-improving the management of early breast cancer: St Gallen international expert consensus on the primary therapy of early breast cancer 2015. Ann Oncol 2015;26:1533–1546.
37. Stahl P, Seeschaaf C, Lebok P, et al.: Heterogeneity of amplification of HER2, EGFR, CCND1, and MYC in gastric cancer. BMC Gastroenterol 2015;15:7.
38. Metzger ML, Behrens HM, Boger C, et al.: MET in gastric cancer —Discarding a 10% cutoff rule. Histopathol 2015.DOI: 10.1111/his.12745
39. Engel KB, Moore HM: Effects of preanalytical variables on the detection of proteins by immunohistochemistry in formalin-fixed, paraffin-embedded tissue. Arch Pathol Lab Med 2011;135: 537–543.
40. Arciero AC: Ki-67 proliferation index and gastric cancer: Answers or more questions. J Surg Oncol 2010;102:199–200.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.