

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/32781> holds various files of this Leiden University dissertation.

Author: Benard, Anne

Title: Epigenetic prognostic biomarkers in colorectal cancer

Issue Date: 2015-04-21

Chapter 5

Prognostic value of Polycomb proteins EZH2, BMI1 and SUZ12 and histone modification H3K27me3 in colorectal cancer

Anne Benard*, Inès J Goossens-Beumer*,
Anneke Q van Hoesel, Hamed Horati,
Hein Putter, Eliane C M Zeestraten,
Cornelis J H van de Velde, Peter J K Kuppen

** Contributed equally*

Plos One, Sept 2014; 9(9):e108265

Abstract

Numerous changes in epigenetic mechanisms have been described in various types of tumors. In search for new biomarkers, we investigated the expression of Polycomb-group (PcG) proteins EZH2, BMI1 and SUZ12 and associated histone modification H3K27me3 in colorectal cancer. PcG proteins and histone modification H3K27me3 were immunohistochemically (IHC) stained on a tissue microarray (TMA), including 247 tumor tissues and 47 normal tissues, and nuclear expression was scored using the semi-automated Ariol system. Tumor tissues showed higher expression of EZH2 ($p=0.05$) and H3K27me3 ($p<0.001$) as compared to their normal counterparts. Combined marker trend analyses indicated that an increase in the number of markers showing high expression was associated with better prognosis. High expression of all four markers in the combined marker analyses was correlated with the best patient survival and the longest recurrence-free survival, with overall survival ($p=0.01$, HR 0.42 (0.21-0.84)), disease-free survival ($p=0.007$, HR 0.23 (0.08-0.67)) and local recurrence-free survival ($p=0.02$, HR 0.30 (0.11-0.84)). In conclusion, we found that expression of PcG proteins and H3K27me3 showed prognostic value in our study cohort. Better stratification of patients was obtained by combining the expression data of the investigated biomarkers as compared to the individual markers, underlining the importance of investigating multiple markers simultaneously.

Introduction

New prognostic biomarkers are warranted in colorectal cancer that could improve decisions for treatment of individual patients in addition to the current TNM (American Joint Committee on Cancer, AJCC (1)) staging system, as even patients with the same TNM classification present with large differences in patient survival and tumor recurrence (2,3). Epigenetic mechanisms have been identified as factors frequently deregulated in tumors and are attractive targets for biomarker research, because of their roles in regulating gene expression and their potentially reversible nature. Numerous changes in DNA methylation, histone modifications and their modifying enzymes have been described in various tumor types, including colorectal cancer (4-6). In this study, we focused on expression of histone-modifying enzymes of the Polycomb-group (PcG) and their associated histone modification, trimethylation of lysine 27 on histone H3 (H3K27me3), in colorectal cancer tissues.

The PcG proteins act in large multi-protein complexes, the so-called Polycomb repressive complexes (PRC) 1 and 2 (7). PcG proteins play an important role in embryonic development and cell proliferation (8,9), and are also involved in inducing epithelial-mesenchymal transition (EMT) (10). Aberrant expression of several PcG proteins and correlations with patient outcome have been reported in various cancers. For example, expression of BMI1 polycomb ring finger oncogene (BMI1), a component of PRC1 and an important factor in stem cells (11,12), was found to be correlated to patient outcome in several types of cancer (13-16). Enhancer of zeste homolog 2 (EZH2), a key protein in the PRC2 complex, was also found to have prognostic value in several types of cancer (17-20). SUZ12 polycomb repressive complex 2 subunit (SUZ12), another key component of the PRC2 complex, was found to have tumor-promoting functions in several cancers, including colon cancer (21-23). The associated histone modification H3K27me3 was found to be higher expressed in tumor tissues, and to be associated with better prognosis in non-small cell lung cancer (24) and breast cancer (19).

Using immunohistochemical staining (IHC) and semi-automated scoring, we studied the expression of PcG proteins EZH2, BMI1 and SUZ12 and their associated histone modification H3K27me3 in a cohort of 247 TNM stage I-III colorectal cancer patients, in correlation with clinical outcome. As the PcG proteins act together on the same histone modification, we hypothesized the combination of all four markers would be more informative with respect to clinical outcome as compared to each of the individual markers.

Materials and methods

Patient selection

Tumor tissues were collected from a consecutive series of 408 colorectal cancer patients who underwent surgical resection of their primary tumor at the Leiden University Medical Center (LUMC) between 1991 and 2001. Patients who underwent pre-operative treatment, who had bilateral tumors, or a history of cancer other than basal cell carcinoma or *in situ* tumors, were excluded from the study analyses. In addition, we included only patients with a histologically proven TNM stage I-III colorectal carcinoma, as determined by an experienced

pathologist. This resulted in a study cohort of 259 patients, with a mean follow-up of 8.6 years. Clinicopathological data were available for all patients in the study cohort. Data were right-censored when patients were alive or free of recurrence at their last follow-up date. Patient characteristics are displayed in Table 1. Patient records information was anonymized and de-identified prior to analysis according to national ethical guidelines (“Code for Proper Secondary Use of Human Tissue”, Dutch Federation of Medical Scientific Societies), and approved by the Medical Ethical Committee of the Leiden University Medical Center (LUMC). This study was performed according to the REMARK guidelines (NCI-EORTC) (25).

Tissue microarray construction and immunohistochemistry

Formalin-fixed paraffin-embedded (FFPE) tumor tissues from each of the patients in the consecutive series of colorectal cancer patients (n=408) were collected from the LUMC pathology archives and used to construct a tissue microarray (TMA), as described previously (26). Sections of 4µm were cut from each TMA block and used for IHC staining. Histologically normal colorectal tissues, as determined by an experienced pathologist, from 47 patients with corresponding tumor tissues included in this study were also collected and prepared for IHC. The following antibodies were used for IHC: anti-EZH2 (612667, BD Biosciences, San Jose, CA, USA), anti-BMI1 (ab14389, Abcam, Cambridge, UK), anti-SUZ12 (ab12073, Abcam) and anti-H3K27me3 (ab6002, Abcam). All antibodies have been validated for use in immunohistochemistry by Western blot (27-30). All primary antibodies were used at predetermined optimal dilutions and IHC was performed using a standard IHC protocol (31). Briefly, endogenous peroxidase was blocked by incubating the sections in a 0.3% solution of hydrogen peroxide (in PBS) for 20 min. Antigen retrieval was performed by heating the sections for 10 min at 95°C in a citrate buffer (pH 6; pH Low Target Retrieval Solution, Dako, Glostrup, Denmark) for EZH2, BMI-1 and H3K27me3 and by heating the sections for 10 min at 95°C in a Tris-EDTA buffer (pH 9; pH High Target Retrieval Solution, Dako) for SUZ12. TMA sections were incubated with the respective primary antibodies overnight (16 hrs). Staining was visualized using the Dako REAL™ EnVision™ Detection System, Peroxidase/DAB+, Rabbit/Mouse (Dako). The stained TMA sections were scanned using a 20x magnification on the semi-automated Ariol system (Leica Microsystems, Wetzlar, Germany). Tumor cell areas (tumor tissues) and colon epithelium (in normal tissues) were marked on the computer screen upon visual inspection, followed by careful training of the Ariol system to correctly identify positively stained and negative nuclei within the marked tissue areas, for each of the markers separately. Nuclear expression, defined as the percentage of positively stained nuclei in the marked area of each tissue core, was then assessed by the Ariol software. Several random cores were evaluated for each TMA section by visual inspection after automated analysis in order to verify correct identification of positively stained nuclei.

Statistical analyses

Data were analyzed in consultation with a statistician (H.P.) using SPSS 20.0 for Windows (SPSS Inc, Chicago, USA). 12 patients were excluded from the statistical analyses, as not all data of all four markers was available for these patients, resulting in a final patient cohort consisted of 247 patients. As the individual marker data were not normally distributed (Shapiro-Wilk test),

TABLE 1. Patient characteristics of the study cohort

		All patients (n=408)		Study cohort (n=247)		p-values
		n	(%)	n	(%)	
Age at randomization						
	<50	45	11	32	13	
	50-75	267	65.4	155	62.8	
	>=75	96	23.6	60	24.3	0.69
Gender						
	Male	202	49.5	127	51.4	
	Female	206	50.5	120	48.6	0.66
TNM stage						
	I	78	19	52	21.1	
	II	149	36.7	110	44.5	
	III	114	27.9	85	34.4	0.21
	IV	67	16.4			
Tumor location						
	Colon	289	71	181	73.3	
	Rectum	119	29	66	26.7	0.6
Tumor size						
	Mean (cm)	4.68		4.71		
	Standard error	2.22		1.53		0.95
MSS status						
	MSS	275	67.2	169	68.4	
	MSI	46	11.2	34	13.8	
	Unknown	87	21.6	44	17.8	0.76
Tumor in follow up						
	No	347	85	209	84.6	
	Yes	61	15	38	15.4	0.91
Adjuvant therapy						
	No	323	79.2	199	80.6	
	Yes	85	20.8	48	19.4	0.97

Patient characteristics are shown for both the complete colorectal cancer series (n=408) and the study cohort (n=247). Patient selection was based on availability of FFPE tissues and available data for all four studied markers. The study cohort selection was representative for the entire colorectal cancer series. P-values represent the results of Student's t-tests. For TNM stage, only tumor stage I-III of the complete patient cohort were compared to the patients in the study cohort.

non-parametric Wilcoxon signed-rank tests were performed to assess the differences in nuclear expression between tumor and paired normal tissues (n=47) for each of the markers. Spearman's signed rank correlation analyses were performed to investigate the correlation between nuclear expression of the individual PcG proteins and histone modification H3K27me3. Cox proportional hazard trend analyses were performed for univariate and multivariate survival analyses of individual markers. Covariates included in all multivariate analyses were age at operation, gender, TNM tumor stage (tumor stages I-III), tumor location, tumor size, microsatellite stability (MSS) status. Covariates "tumor in the follow up" and "adjuvant therapy" were entered as time-dependent covariates. Overall survival (OS) was defined as the time from surgery until death (by any cause). Disease-free survival (DFS) was defined as the time from surgery until the occurrence of a second primary colorectal tumor, locoregional recurrence or distant recurrence, or death by colorectal cancer. Locoregional recurrence-free survival (LRRFS) was defined as the time from surgery until the occurrence of a locoregional recurrence or death by cancer. Distant recurrence-free survival (DRFS) was defined as the time from surgery until the occurrence of a distant recurrence or death by cancer.

On the basis of the skewed distribution of expression data of each of the individual markers, the median expression was used to divide the patients into high expression (above-median) and low expression (below-median) groups. The four markers were then combined into a new variable, based on the number of markers showing high nuclear expression, resulting in the following grouping: all low (group 1), 1 high (group 2), 2 high (group 3), 3 high (group 4) and all high (group 5). Univariate and multivariate Cox proportional hazard analyses were performed using the group numbers as a categorical variable, using group 1 (all low) as the reference group. On the basis of these results we decided to combine patient groups 2, 3, and 4 into one patient group; all further statistical analyses were performed using three patient groups. In addition to the Cox proportional hazard analyses, trend analyses were performed using the group numbers as continuous variables to assess the influence of the combined markers on patient survival and tumor recurrence. Resulting hazard ratios (HR) represent the HR for each unit of increase (increase in group number). Cumulative incidence curves were made for DFS, LRRFS and DRFS, accounting for competing risks (32). Kaplan-Meier curves were used to visualize differences between the three patient groups for OS. For all statistical analyses, two-sided p-values ≤ 0.05 were considered statistically significant, and p-values $0.05 < p \leq 0.1$ were considered a trend.

Results

Expression in tumor versus paired normal colorectal tissues

Nuclear expression of all individual markers (EZH2, BMI1, SUZ12 and H3K27me3) in tumor tissues was compared to nuclear expression in paired normal colorectal tissues. When analyzing expression differences in the study cohort as a whole, only median H3K27me3 and EZH2 expression were significantly different between tumor and normal tissues ($p < 0.001$ and $p = 0.05$, respectively; Figure 1A). In individual tumors, however, all markers showed marked differences in expression compared to their normal counterparts (Figure 1B). Survival analyses based on below- or above-median expression in the normal tissues did not show differences in patient survival or tumor recurrence (data not shown).

Figure 1

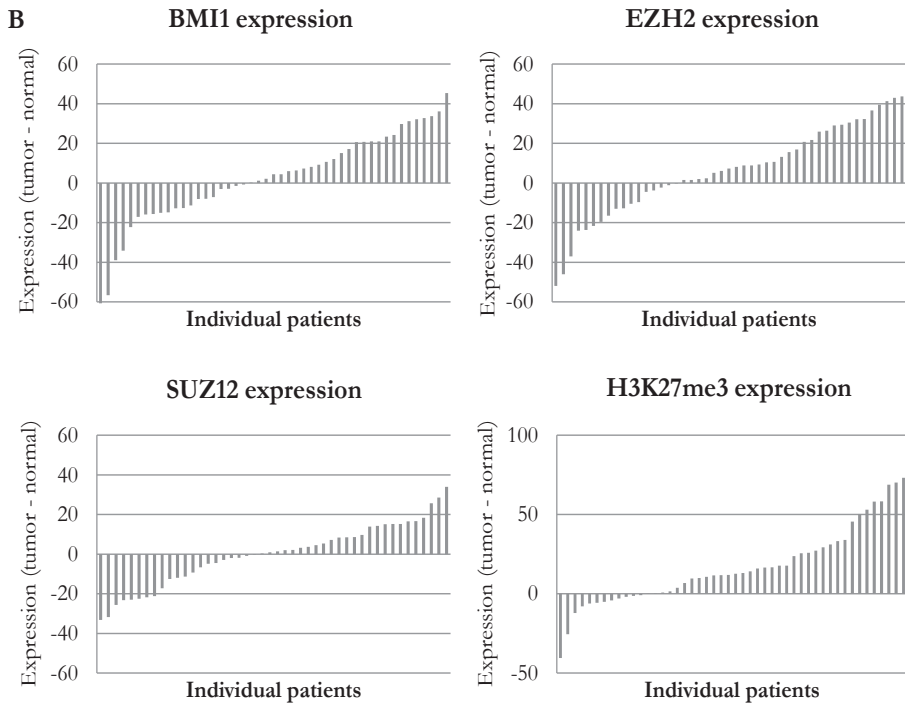
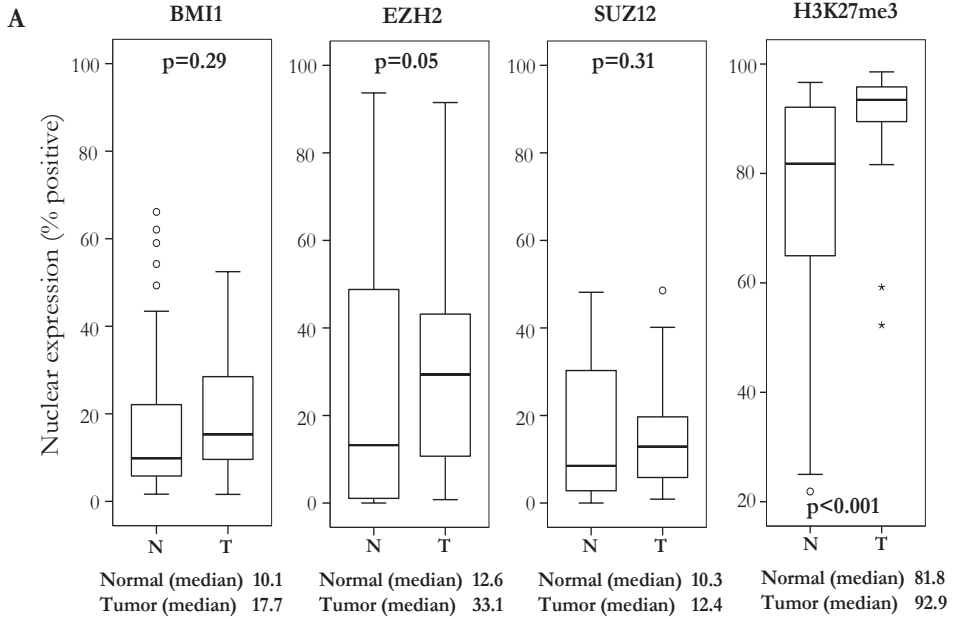


Figure legends: see next page

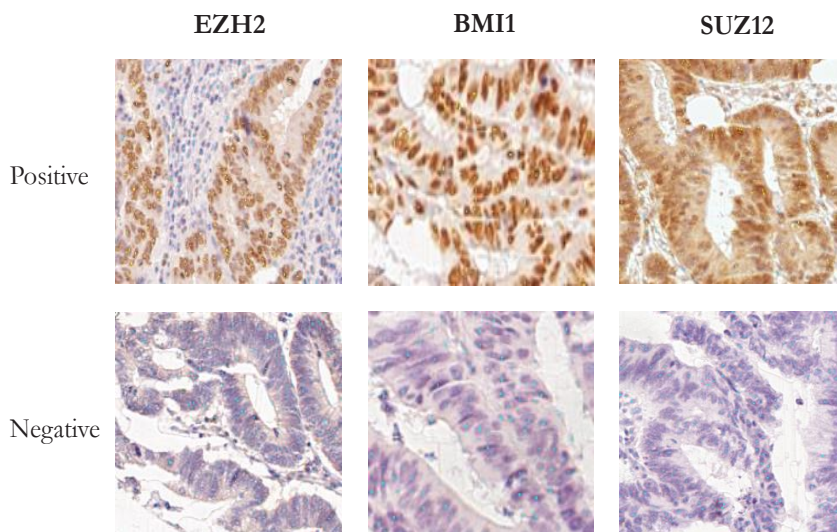
5

Figure 1 (see previous page) **A.** Differences in nuclear expression between normal and tumor tissues of individual markers. Displayed are differences in nuclear expression, measured as the percentage of positively stained nuclei, between normal and tumor tissues ($n=47$). Boxplots show the median and range of expression of each of the individual markers in normal (N) and tumor (T) samples. The median percentages of positive nuclei are given for each of the markers. P-values represent statistical differences between normal and tumor samples, calculated using the Wilcoxon signed rank test. **B.** Histograms show the difference in expression between tumor and paired normal tissues (y-axis) for each of the individual patients (x-axis). Differences in expression were calculated as follows: expression difference = expression in tumor tissue – expression in normal tissue. Negative values indicate higher expression in normal tissues, positive values indicate higher expression in tumor tissues.

Individual marker analyses in tumor tissues

Examples of identification of positive and negative tumor cell nuclei for each of the individual markers by the Ariol system are shown in Figure 2. We first analyzed if the expression of histone modification H3K27me3 was correlated to the expression of the individual PcG proteins. The nuclear expression (percentage of positive nuclei) of H3K27me3 was indeed positively correlated with expression of EZH2 ($p<0.001$), BMI1 ($p<0.001$) and SUZ12 ($p=0.05$). No correlation was observed between the expression of the individual markers and TNM tumor stage. For survival analyses, patients were divided into low and high expression groups on the basis of the median expression of each of the individual markers, as given in Figure 1A. In survival analyses of individual markers, BMI1 showed strong correlations to patient survival (OS and DFS) and tumor recurrence (LRRFS and DRFS) in both univariate and multivariate analyses (Table 2). EZH2 and H3K27me3 showed significant correlations for DFS only.

Figure 2



Identification of positive and negative tumor cell nuclei by the Ariol system. The Ariol system trainer overlay shows correct identification of positive (indicated by yellow dots) and negative (blue dots) nuclei in tumor cores. TMA slides were scanned using a 20x magnification. Shown for all markers are positively stained tumor cores (top row) and negative tumor cores (bottom row). The Ariol system was trained to identify positive and negative cells for each marker individually.

TABLE 2. Univariate and multivariate survival analyses individual markers

			EZH2	BMI1	SUZ12	H3K27me3
OS	Univariate	p-value	<i>0.07</i>	0.05	0.9	0.5
		HR	0.74	0.73	1.03	0.89
		(95% CI)	(0.54-1.03)	(0.53-1.00)	(0.73-1.47)	(0.64-1.24)
	Multivariate	p-value	0.3	0.009	0.3	0.5
		HR	0.84	0.62	0.83	0.88
		(95% CI)	(0.60-1.18)	(0.44-0.89)	(0.57-1.20)	(0.62-1.24)
DFS	Univariate	p-value	0.04	<i>0.08</i>	0.8	<i>0.07</i>
		HR	0.64	0.68	0.95	0.66
		(95% CI)	(0.42-0.99)	(0.44-1.04)	(0.60-1.49)	(0.42-1.03)
	Multivariate	p-value	<i>0.08</i>	0.03	0.2	0.05
		HR	0.67	0.61	0.71	0.64
		(95% CI)	(0.43-1.05)	(0.39-0.96)	(0.44-1.17)	(0.41-0.99)
LRRFS	Univariate	p-value	<i>0.06</i>	0.03	0.8	0.2
		HR	0.67	0.63	1.06	0.76
		(95% CI)	(0.44-1.03)	(0.41-0.96)	(0.67-1.68)	(0.49-1.18)
	Multivariate	p-value	<i>0.1</i>	0.005	0.6	0.2
		HR	0.71	0.52	0.88	0.76
		(95% CI)	(0.46-1.11)	(0.33-0.82)	(0.55-1.42)	(0.49-1.19)
DRFS	Univariate	p-value	0.04	<i>0.06</i>	0.6	0.4
		HR	0.64	0.66	1.11	0.83
		(95% CI)	(0.42-0.99)	(0.48-1.02)	(0.69-1.77)	(0.53-1.29)
	Multivariate	p-value	<i>0.1</i>	0.02	0.7	0.6
		HR	0.73	0.58	0.91	0.89
		(95% CI)	(0.46-1.15)	(0.36-0.92)	(0.56-1.48)	(0.57-1.42)

Shown are the results of the univariate and multivariate analyses of all individual markers, with all p-values and hazard ratios (HR) and their 95% confidence intervals (95% CI). The "low expression" group was used as the reference group. OS = overall survival, DFS = disease-specific survival, LRRFS = locoregional recurrence-free survival, DRFS = distant recurrence-free survival. Significant p-values are indicated in **bold**, p-values showing a trend (between 0.5 and 1.0) in *italic*.

For all three markers (BMI1, EZH2 and H3K27me3), high expression was associated with better patient survival as compared to the patients showing low expression, with p-values for DFS of p=0.07 (BMI1), p=0.04 (EZH2) and p=0.06 (H3K27me3). SUZ12 did not show any differences in patient survival or tumor recurrence based on low or high expression of the marker in tumor tissues.

Combined markers

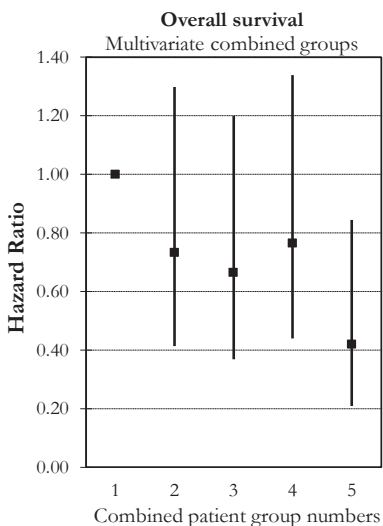
We hypothesized that combining multiple markers would result in better stratification of patients. Therefore, we performed statistical analyses on combinations of the histone-modifying enzymes. These analyses showed that combining multiple markers indeed resulted in statistically significant differences between the patient groups and more pronounced hazard ratios, indicating a more pronounced effect on patient survival. The combination of histone-modifying enzymes EZH2 and BMI1 showed significant differences for both patient survival and recurrence-free survival, with $p=0.02$ (HR=0.72; 95% CI 0.54-0.94) for DFS and $p=0.012$ (HR=0.71; 95% CI 0.54-0.92) for LRRFS in multivariate analyses. Combining EZH2 and SUZ12 showed a trend for DFS in multivariate analyses, with $p=0.08$ (HR=0.77; 95% CI 0.57-1.04). The combination of BMI1 and SUZ12 showed significant differences for patient survival, with $p=0.02$ (HR=0.76; 95% CI 0.61-0.96) for overall survival and $p=0.05$ (HR=0.73; 95% CI 0.54-1.00) for DFS.

Figure 3

A Trend analysis combined markers

EZH2, BMI1, SUZ12 and H3K27me3	OS			DSS			LRRFS			DRFS		
	p-value	HR	(95% CI)	p-value	HR	(95% CI)	p-value	HR	(95% CI)	p-value	HR	(95% CI)
Univariate	0.09	0.88	(0.77-1.02)	0.04	0.83	(0.69-0.99)	0.08	0.85	(0.71-1.02)	0.12	0.86	(0.72-1.04)
Multivariate	0.05	0.87	(0.76-1.00)	0.01	0.79	(0.65-0.95)	0.04	0.83	(0.69-0.99)	0.12	0.86	(0.71-1.04)

B



C

Overall survival combined markers

Univariate	p-value	HR	(95% CI)
Group 1	0.07		
Group 2	0.02	0.53	(0.31-0.92)
Group 3	0.02	0.52	(0.29-0.90)
Group 4	0.07	0.62	(0.36-1.04)
Group 5	0.01	0.41	(0.21-0.80)

Multivariate	p-value	HR	(95% CI)
Group 1	0.2		
Group 2	0.3	0.73	(0.41-1.29)
Group 3	0.2	0.67	(0.37-1.20)
Group 4	0.3	0.77	(0.44-1.34)
Group 5	0.02	0.42	(0.21-0.84)

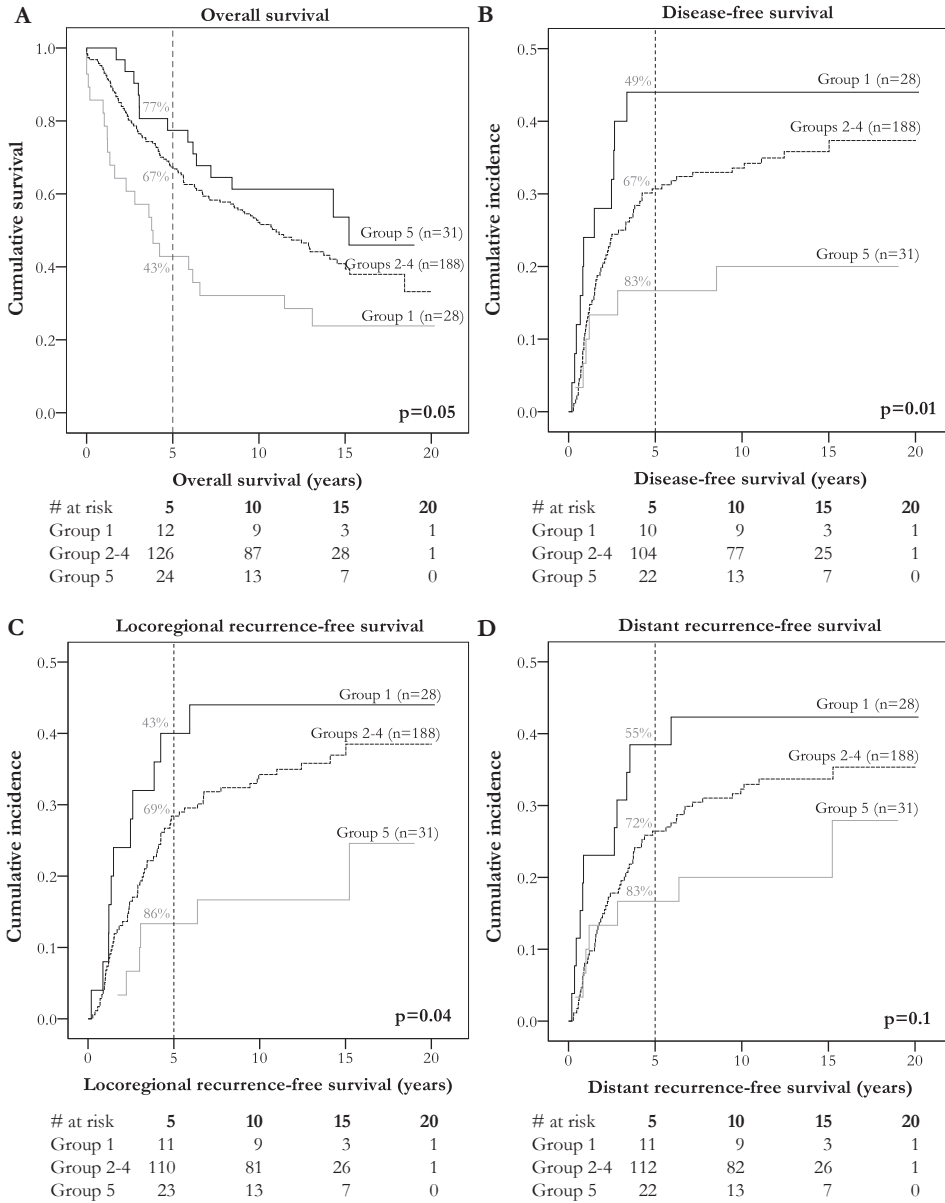
Trend analyses and hazard ratios of combined marker groups. Patient groups were made on the basis of the number of markers showing high expression: all low (group 1), one high (group 2), two high (group 3), three high (group 4) and all high (group 5). **A.** Univariate and multivariate Cox regression trend analyses were performed using the combined marker groups as continuous variables. Hazard ratios per unit increase of each of the patient groups were plotted (**B**) and listed (**C**) for both the univariate and multivariate Cox regression trend analyses using the combined marker groups as categorical variables. The numbers of patients in the individual patient groups were: group 1 (n=28), group 2 (n=59), group 3 (n=55), group 4 (n=74) and group 5 (n=31).

Because the three PcG proteins act together in multi-protein complexes to regulate H3K27me3 expression, we hypothesized that combining all markers (BMI1, EZH2, SUZ12 and H3K27me3) into one variable would result in even better stratification of patients. Patients were divided into five groups on the basis of the number of markers showing high nuclear expression. This resulted in the following patient groups: all low (group 1), one high (group 2), two high (group 3), three high (group 4) and all high (group 5). Patient characteristics of the 5 patient groups were comparable to the study cohort (Table S1). Multivariate trend analyses of the combined markers, using the patient group numbers as continuous variables, showed overall hazard ratios of 0.79-0.88 for each additional marker showing high nuclear expression in both univariate and multivariate analyses, indicating better patient survival and lower chances of tumor recurrence for each additional marker showing high expression (Figure 3A). When the patient group numbers were entered as categorical variables, a similar trend was observed (Figures 3B and 3C). Generally, hazard ratios for OS decreased with increasing group number, indicating a better patient survival when more markers were highly expressed compared to the “all low” expression group (group 1). Patients showing high expression of all markers, the “all high” group (group 5), showed the best overall survival ($p=0.01$, HR=0.42, 95% CI 0.21-0.84) as compared to reference group 1, which showed the shortest survival. Groups 2, 3 and 4 showed similar hazard ratios (Figures 3B and 3C), which was also reflected in the Kaplan-Meier survival curves that run close together as compared to the survival curves of groups 1 and 5. Therefore, we decided to combine the three patient groups 2,3 and 4 into one group, resulting in three patient groups (group 1, groups 2-4 and group 5). Kaplan-Meier curves and cumulative incidence plots showed significant differences between the three resulting patient groups for OS, DFS, LRRFS and a trend for DRFS, which were also reflected in the 5-year survival rates (Figure 4). The best patient survival and longest recurrence-free periods were observed for patients showing high expression of all four markers (“all high”, group 5) in the tumor samples, with 5-year survival rates of 77% for OS, 83% for DFS and DRFS, and 86% for LRFS. Patients in combined groups 2-4 showed shorter OS, DFS and LRFS, with 5-year survival rates of 67% for OS and DFS, 69% for LRFS and 72% for DRFS. Patients in the “all low” group (group 1) showed significantly shorter OS, DFS and LRFS compared to either of the other patient groups, with 5-year survival rates of 43% for OS and LRFS, 49% for DFS and 55% for DRFS. Taken together, 5-year survival rates were lower when more markers showed low expression. The hazard ratios in both univariate and multivariate also reflected these findings (Table 3): group 5 shows the lowest hazard ratio as compared to reference group 1 (for example, multivariate HR=0.23 (0.08-0.67) for DFS). This indicates a lower risk of an event (patient death or locoregional tumor recurrence) for patients in the “all high” group for OS, DFS and LRFS. For DRFS, statistically significant results were only observed in univariate analyses.

Discussion

In addition to gene mutations, aberrant expression patterns of epigenetic regulators have been recognized as crucial events in the tumorigenic process, resulting in marked changes in gene expression. Changes in the expression of these epigenetic regulators include DNA methyltransferases and consequent changes in DNA methylation profiles, and histone-modifying enzymes and resulting changes in their corresponding histone modifications. In this study, we

Figure 4



Survival curves for the combined marker groups. Combined marker (EZH2, BMI1, SUZ12 and H3K27me3) expression groups were divided into three patient groups: group 1, groups 2-4 and group 5. The numbers of patients in the individual patient groups were: group 1 (n=28), group 2 (n=59), group 3 (n=55), group 4 (n=74) and group 5 (n=31). Kaplan-Meier curves were made for overall survival (A) and cumulative incidence curves are shown for disease-free survival (B), locoregional recurrence-free survival (C) and distant recurrence-free survival (D). 5-year survival rates are given for each patient group. Tables below the curves indicate the numbers at risk (#) per group for the different time points.

TABLE 3. Univariate and multivariate analyses combined markers

		Univariate			Multivariate		
		p-value	HR	(95% CI)	p-value	HR	(95% CI)
OS	Group 1	0.018			0.05		
	Groups 2-4	0.014	0.56	(0.35-0.89)	0.21	0.72	(0.44-1.19)
	Group 5	0.009	0.41	(0.21-0.80)	0.01	0.42	(0.21-0.84)
DFS	Group 1	<i>0.07</i>			0.03		
	Groups 2-4	0.1	0.64	(0.33-1.21)	0.2	0.61	(0.31-1.22)
	Group 5	0.02	0.32	(0.12-0.86)	0.007	0.23	(0.08-0.67)
LRRFS	Group 1	<i>0.06</i>			0.03		
	Groups 2-4	0.2	0.64	(0.34-1.22)	0.96	1.02	(0.49-2.09)
	Group 5	0.02	0.31	(0.11-0.83)	0.02	0.30	(0.11-0.84)
DRFS	Group 1	0.12			0.24		
	Groups 2-4	0.11	0.59	(0.31-1.13)	0.66	0.85	(0.41-1.75)
	Group 5	0.05	0.38	(0.15-0.98)	0.11	0.44	(0.16-1.20)

Results of Cox proportional hazard univariate and multivariate analyses are shown for combined markers EZH2, BMI1, SUZ12 and H3K27me3, with p-values and hazard ratios with their 95% confidence intervals. Patient groups were made based on the number of markers showing high expression: all low (group 1), one, two or three high (groups 2-4) and all high (group 5). Group 1 was used as the reference group. Covariates included in all multivariate analyses were age at operation, gender, TNM tumor stage, tumor location, tumor size, microsatellite stability (MSS) status, tumor in the follow up and adjuvant therapy. Significant values are shown in **bold**, p-values showing a trend (between 0.5 and 1.0) in *Italic*.

investigated the expression of three PcG proteins (EZH2, BMI1 and SUZ12) and associated histone modification H3K27me3 in colorectal cancer tissues. Aberrant expression of each of these histone-modifying enzymes, and of histone modification H3K27me3, has been indicated to contribute to tumorigenesis in several types of cancer and has been correlated to patient outcome (11-24). Studies in literature show conflicting results regarding the prognostic value of the Polycomb-group proteins in colorectal cancer. For example, high EZH2 expression has been associated with poor prognosis in a series of colorectal cancer patients by Wang *et al.* (33), whereas high EZH2 expression was found to be associated with better relapse-free survival in colon cancer patients (but not in rectal cancer patients) by Fluge *et al.* (34). In addition, high expression of BMI1 was found to correlate with good prognosis in breast cancer in a study by Pietersen *et al.* (35), whereas high BMI1 was associated with poor prognosis in colon cancer in a study by Du *et al.* (36). In our study cohort, survival data for the individual markers showed that high expression of all markers was correlated with better patient survival and longer recurrence-free periods as compared to patients showing low expression. The results found in this study correspond to our previous findings that high expression of H3K27me3 was associated with better patient survival in rectal tumors (4). In this study, we showed that high expression of H3K27me3 was indeed associated with better patient survival and longer recurrence-free

periods. As we showed that the expression of the PcG proteins was directly related to the expression of H3K27me₃, we expected a similar correlation of expression of the PcG proteins with clinical outcome, which was indeed confirmed by the results presented in this manuscript. High levels of H3K27me₃, because of aberrant expression of PcG proteins, might prevent aberrant expression of oncogenes, activation of retrotransposon sequences (such as LINE-1; 4), and result in other (epi)genomic events that promote tumor aggressiveness.

In addition to the individual markers, combinations of PcG proteins in correlation with patient outcome have been studied by several research groups. For example, co-expression of EZH2 and BMI1 was reported to be associated with poor prognosis in various cancers (37-39). In contrast, overexpression of EZH2 and BMI1 were reported to have different influences on patient prognosis in breast cancer (35), and was found to have no prognostic value in urothelial carcinoma of the bladder (40). In our colorectal cancer study cohort, all combinations of histone-modifying enzymes showed prognostic value. In order to obtain more information about epigenetic pathways with potential prognostic value in colorectal cancer, we performed multivariate survival analyses using combined expression data of multiple PcG proteins (EZH2, BMI1 and SUZ12) and their associated histone modification H3K27me₃. Combining the three PcG proteins and their associated histone modification resulted in significantly better stratification of patient groups as compared to the individual markers. In combined marker analyses, the best patient survival and longest recurrence-free periods were observed for patients showing high expression of all four markers (“all high”) in the tumor samples. Patients in the “all low” group showed significantly shorter OS, DFS and LRFS compared to either of the other patient groups. The results of the combined marker analyses underline the co-operation of these three enzymes in PcG complexes, and thus provide a better risk stratification of patients.

In addition to the roles of EZH2, BMI1 and SUZ12 in epigenetic regulation of chromatin structure and gene expression, direct regulation of protein function has been described for EZH2 and BMI1, including protein phosphorylation and ubiquitination. A cytosolic EZH2 and SUZ12-containing methyltransferase complex has been linked to actin polymerization, an important process in cell proliferation (41). Shuttling of the EZH2 and SUZ12 containing complex between different cellular compartments may explain the weak cytosolic staining observed for EZH2 and SUZ12 in addition to the strong nuclear staining for these markers, as compared to the strict nuclear staining observed for BMI1. Another non-histone protein methylated by EZH2 is cardiac transcription factor GATA4. Methylation reduces its transcriptional activity, resulting in inhibition of proper cardiac development (42). These examples indicate that aberrant expression of these PcG proteins influences key processes such as gene transcription and cell proliferation, promoting the transformation of normal cells into tumor cells.

In conclusion, we showed that combined expression of PcG proteins EZH2, BMI1 and SUZ12 and their associated histone modification H3K27me₃ has prognostic value in our colorectal cancer study cohort. Combined marker expression resulted in better stratification of patients as compared to the individual markers and hence provides more insight into the roles of these epigenetic proteins and –modifications in colorectal cancer. Other combinations of epigenetic mechanisms should be investigated in colorectal cancer to further unravel the underlying biology in individual tumors. This will advance the search for new biomarkers to be used in a clinical setting in order to better classify patients for treatment.

Acknowledgements

We would like to thank C.M. Janssen for providing us with the tissue microarray sections used for immunohistochemistry.

References

- (1) Edge SB. AJCC Cancer Staging Manual, 7th edition. 7th ed. Springer: New York, 2010.
- (2) Hoeben KW, van Steenberghe LN, van de Wouw AJ *et al.* Treatment and complications in elderly stage III colon cancer patients in the Netherlands. *Ann Oncol* 2013; **24**(4):974-9.
- (3) van den Broek CB, Bastiaannet E, Dekker JW *et al.* Time trends in chemotherapy (administration and costs) and relative survival in stage III colon cancer patients - a large population-based study from 1990 to 2008. *Acta Oncol* 2013; **52**(5):941-9.
- (4) Benard A, van de Velde CJ, Lessard L *et al.* Epigenetic status of LINE-1 predicts clinical outcome in early-stage rectal cancer. *Br J Cancer* 2013; **109**(12):3073-83.
- (5) de Maat MF, van de Velde CJ, Benard A *et al.* Identification of a quantitative MINT locus methylation profile predicting local regional recurrence of rectal cancer. *Clin Cancer Res* 2010; **16**(10):2811-8.
- (6) van Engeland M., Derks S, Smits KM *et al.* Colorectal cancer epigenetics: complex simplicity. *J Clin Oncol* 2011; **29**(10):1382-91.
- (7) Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* 2009; **10**(10):697-708.
- (8) Aloia L., Di SB, Di CL. Polycomb complexes in stem cells and embryonic development. *Development* 2013; **140**(12):2525-34.
- (9) Oktaba K, Gutierrez L, Gagneur J *et al.* Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. *Dev Cell* 2008; **15**(6):877-89.
- (10) Malouf GG, Taube JH, Lu Y *et al.* Architecture of epigenetic reprogramming following Twist1 mediated epithelial-mesenchymal transition. *Genome Biol* 2013; **14**(12):R144.
- (11) Manoranjan B, Wang X, Hallett RM *et al.* FoxG1 interacts with Bmi1 to regulate self-renewal and tumorigenicity of medulloblastoma stem cells. *Stem Cells* 2013; **31**(7):1266-77.
- (12) Siddique HR, Parray A, Zhong W *et al.* BMI1, stem cell factor acting as novel serum-biomarker for Caucasian and African-American prostate cancer. *PLoS One* 2013; **8**(1):e52993.
- (13) Allegra E, Puzzo L, Zuccala V *et al.* Nuclear BMI-1 expression in laryngeal carcinoma correlates with lymph node pathological status. *World J Surg Oncol* 2012; **10**:206.
- (14) Gialmanidis IP, Bravou V, Petrou I *et al.* Expression of Bmi1, FoxF1, Nanog, and gamma-catenin in relation to hedgehog signaling pathway in human non-small-cell lung cancer. *Lung* 2013; **191**(5):511-21.
- (15) Ruan ZP, Xu R, Lv Y *et al.* Bmi1 knockdown inhibits hepatocarcinogenesis. *Int J Oncol* 2013; **42**(1):261-8.
- (16) Yoshikawa R, Tsujimura T, Tao L *et al.* The oncoprotein and stem cell renewal factor BMI1 associates with poor clinical outcome in oesophageal cancer patients undergoing preoperative chemoradiotherapy. *BMC Cancer* 2012; **12**:461.
- (17) Behrens C, Solis LM, Lin H *et al.* EZH2 protein expression associates with the early pathogenesis, tumor progression, and prognosis of non-small cell lung carcinoma. *Clin Cancer Res* 2013; **19**(23):6556-65.
- (18) Crea F, Fornaro L, Paolicchi E *et al.* An EZH2 polymorphism is associated with clinical outcome in metastatic colorectal cancer patients. *Ann Oncol* 2012; **23**(5):1207-13.
- (19) Holm K, Grabau D, Lovgren K *et al.* Global H3K27 trimethylation and EZH2 abundance in breast tumor subtypes. *Mol Oncol* 2012; **6**(5):494-506.
- (20) Mu Z, Li H, Fernandez SV *et al.* EZH2 knockdown suppresses the growth and invasion of human inflammatory breast cancer cells. *J Exp Clin Cancer Res* 2013; **32**(1):70.
- (21) Benoit YD, Witherspoon MS, Laursen KB *et al.* Pharmacological inhibition of polycomb repressive complex-2 activity induces apoptosis in human colon cancer stem cells. *Exp Cell Res* 2013; **319**(10):1463-70.
- (22) Cui Y, Chen J, He Z *et al.* SUZ12 depletion suppresses the proliferation of gastric cancer cells. *Cell Physiol Biochem* 2013; **31**(6):778-84.
- (23) Li H, Cai Q, Wu H *et al.* SUZ12 promotes human epithelial ovarian cancer by suppressing apoptosis via silencing HRK. *Mol Cancer Res* 2012; **10**(11):1462-72.
- (24) Chen X, Song N, Matsumoto K *et al.* High expression of trimethylated histone H3 at lysine 27 predicts better prognosis in non-small cell lung cancer. *Int J Oncol* 2013; **43**(5):1467-80.
- (25) McShane LM, Altman DG, Sauerbrei W *et al.* Reporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* 2005; **93**(4):387-91.
- (26) Zeestraten EC, Reimers MS, Saadatmand S *et al.* Combined analysis of HLA class I, HLA-E and HLA-G predicts prognosis in colon cancer patients. *Br J Cancer* 2013.
- (27) Bender S, Tang Y, Lindroth AM *et al.* Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell* 2013; **24**(5):660-72.

- (28) Hosogane M, Funayama R, Nishida Y *et al.* Ras-induced changes in H3K27me3 occur after those in transcriptional activity. *PLoS Genet* 2013; **9**(8):e1003698.
- (29) Maertens GN, El Messaoudi-Aubert S, Racek T *et al.* Several distinct polycomb complexes regulate and co-localize on the INK4a tumor suppressor locus. *PLoS One* 2009; **4**(7):e6380.
- (30) Tong Q, He S, Xie F *et al.* Ezh2 Regulates Transcriptional and Posttranslational Expression of T-bet and Promotes Th1 Cell Responses Mediating Aplastic Anemia in Mice. *J Immunol* 2014; **192**(11):5012-22.
- (31) van Nes JG, de Kruijf EM, Faratian D *et al.* COX2 expression in prognosis and in prediction to endocrine therapy in early breast cancer patients. *Breast Cancer Res Treat* 2011; **125**(3):671-85.
- (32) Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; **26**(11):2389-430.
- (33) Wang CG, Ye YJ, Yuan J *et al.* EZH2 and STAT6 expression profiles are correlated with colorectal cancer stage and prognosis. *World J Gastroenterol* 2010; **16**(19):2421-7.
- (34) Fluge O, Gravdal K, Carlsen E *et al.* Expression of EZH2 and Ki-67 in colorectal cancer and associations with treatment response and prognosis. *Br J Cancer* 2009; **101**(8):1282-9.
- (35) Pietersen AM, Horlings HM, Hauptmann M *et al.* EZH2 and BMI1 inversely correlate with prognosis and TP53 mutation in breast cancer. *Breast Cancer Res* 2008; **10**(6):R109.
- (36) Du J, Li Y, Li J *et al.* Polycomb group protein Bmi1 expression in colon cancers predicts the survival. *Med Oncol* 2010; **27**(4):1273-6.
- (37) Yonemitsu Y, Imazeki F, Chiba T *et al.* Distinct expression of polycomb group proteins EZH2 and BMI1 in hepatocellular carcinoma. *Hum Pathol* 2009; **40**(9):1304-11.
- (38) Wu Z, Wang Q, Wang L *et al.* Combined aberrant expression of Bmi1 and EZH2 is predictive of poor prognosis in glioma patients. *J Neurol Sci* 2013; **335**(1-2):191-6.
- (39) Ha SY, Kim SH. Co-expression of Bmi1 and EZH2 as an independent poor prognostic factor in esophageal squamous cell carcinoma. *Pathol Res Pract* 2012; **208**(8):462-9.
- (40) Hinz S, Kempkensteffen C, Christoph F *et al.* Expression parameters of the polycomb group proteins BMI1, SUZ12, RING1 and CBX7 in urothelial carcinoma of the bladder and their prognostic relevance. *Tumour Biol* 2008; **29**(5):323-9.
- (41) Su IH, Dobenecker MW, Dickinson E *et al.* Polycomb group protein ezh2 controls actin polymerization and cell signaling. *Cell* 2005; **121**(3):425-36.
- (42) He A, Shen X, Ma Q *et al.* PRC2 directly methylates GATA4 and represses its transcriptional activity. *Genes Dev* 2012; **26**(1):37-42.

Supplementary TABLE 1 Patient characteristics of all patient groups used in combined-marker analyses

	Group 1 (n=28)		Group 2 (n=59)		Group 3 (n=55)		Group 4 (n=74)		Group 5 (n=31)		p-values	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)		
Age at randomization												
	<50	1	3.6	9	15.3	9	16.4	10	13.5	3	9.7	0.81
	50-75	15	53.6	41	69.5	36	65.5	47	63.5	16	51.6	
	>=75	12	42.9	9	15.3	10	18.2	17	23	12	38.7	
Gender												
	Male	14	50	28	47.5	27	49.1	36	48.6	15	48.4	0.99
	Female	14	50	31	52.5	28	50.9	38	51.4	16	51.6	
TNM stage												
	I	7	25	12	20.3	12	21.8	14	18.9	7	22.6	0.56
	II	9	32.1	27	45.8	25	45.5	31	41.9	18	58.1	
	III	12	42.9	20	33.9	18	32.7	29	39.2	6	19.4	
Tumor location												
	Colon	18	64.3	44	74.6	40	72.7	61	82.4	18	58.1	0.85
	Rectum	10	35.7	15	25.4	15	27.3	13	17.6	13	41.9	
Tumor size												
	Mean (cm)	4.55		4.5		4.72		4.89		4.77		0.91
	Standard error	0.55		0.31		0.32		0.37		0.41		
MSS status												
	MSS	18	64.2	44	74.6	35	63.6	51	68.9	21	67.7	0.73
	MSI	5	17.9	7	11.9	10	18.2	8	10.8	4	12.9	
	Unknown	5	17.9	8	13.6	10	18.2	15	20.3	6	19.4	
Tumor in follow up*												
	No	25	89.3	52	88.1	45	81.8	65	87.8	22	71	0.52
	Yes	3	10.7	7	11.9	10	18.2	9	12.2	9	29	
Adjuvant therapy												
	No	25	89.3	45	76.3	48	87.3	55	74.3	26	83.9	0.13
	Yes	3	10.7	14	23.7	7	12.7	19	25.7	5	16.1	

Patient characteristics are shown for all patient groups as used in the combined-marker analyses. The patient groups show comparable patient characteristics to the complete study cohort of 247 patients (Table 1). P-values represent the Jonckheere-Terpstra test used to test if samples came from the same distribution. For the variable "tumor size", a one-way ANOVA test was performed to test for statistical differences between the patient groups. * = second primary tumor during follow-up period. MSS = microsatellite stable, MSI = microsatellite instable.

