uicc
global cancer control

IJC
International Journal of Cancer

Lymph node CEA and MUC2 mRNA as useful predictors of outcome in colorectal cancer

Lina Ohlsson[1], Anne Israelsson[1], Åke Öberg[2], Richard Palmqvist[3], Hans Stenlund[4], Marie-Louise Hammarström[1], Sten Hammarström[1] and Gudrun Lindmark[5]

[1] Department of Clinical Microbiology, Immunology, Umeå University, Umeå, Sweden
[2] Department of Surgery and Perioperative Sciences, Surgery, Umeå University, Umeå, Sweden
[3] Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden
[4] Department of Public Health and Clinical Medicine, Epidemiology and Global Health, Umeå University, Umeå, Sweden
[5] Department of Surgery, Helsingborg Hospital/Lund University, Helsingborg, Sweden

The aim was to explore the utility for staging and prognostic impact of carcinoembryonic antigen (CEA), cytokeratin 20 (CK20), guanylyl cyclase C (GCC), CUB (complement protein subcomponents C1r/C1s, urchin embryonic growth factor, and bone morphogenic protein 1) containing domain protein 1 (CDCP1) and mucin 2 (MUC2) mRNA levels in mesenteric lymph nodes of colorectal cancer (CRC) patients. Lymph nodes were collected at surgery and bisected; one half was subjected to biomarker mRNA analysis using real-time quantitative RT-PCR and the other half to routine histopathology. Lymph nodes from 174 CRC patients and 24 controls were analyzed. The median follow-up time was 59 (range 17–131) months. Cut-off levels were defined by analyzing quintiles by Cox regression model. CEA mRNA showed the best discriminating power between patients with recurrence in CRC after surgery and patients who were apparently disease-free ($p = 0.015$). The risk of recurrence for the CEA(+) patients was 4.6 times greater than for the CEA(−) patients ($p < 0.0001$). The other biomarkers gave lower hazard ratios. Cumulative survival analysis demonstrated that the average survival time was 99 months for CEA(−) patients compared to 39 months for CEA(+) patients, a difference of 60 months ($p < 0.0001$). Six to nine percent of the Stage I and Stage II patients [H&E(−)] had CEA(+), CK20(+), GCC(+) and/or MUC2(+) lymph nodes. Two of these patients died from recurrent CRC. Low lymph node MUC2/CEA mRNA ratio identified patients with high risk for recurrence ($p = 0.011$). Thus, quantitative reverse transcriptase-polymerase chain reaction of CEA mRNA is a sensitive method to identify tumor cells in lymph nodes of CRC patients and, in combination with MUC2 mRNA, allows improved prediction of clinical outcome.

Early detection of colorectal cancer (CRC) in the localized tumor stage greatly improves the patient's chance of survival.[1]

Key words: colorectal neoplasms, lymphatic metastasis, neoplasm staging, prognosis, qRT-PCR

Abbreviations: CDCP1: CUB (complement protein subcomponents C1r/C1s, urchin embryonic growth factor, and bone morphogenic protein 1) containing domain protein 1 (CDCP1); CEA: carcinoembryonic antigen; CK20: cytokeratin 20; CRC: colorectal cancer; EC: epithelial cell; GCC: guanylyl cyclase C; H&E: hematoxylin and eosin; MUC2: mucin 2; PBMC: peripheral blood mononuclear cell; qRT-PCR: quantitative reverse transcriptase-polymerase chain reaction; UC: ulcerative colitis

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: Stig and Ragna Gorthon's Foundation in Helsingborg, Gunnar Nilsson's Cancer Foundation, Swedish Cancer Foundation, Norrland's Cancer Foundation, County of Västerbotten
DOI: 10.1002/ijc.26182
History: Received 16 Dec 2010; Accepted 19 Apr 2011; Online 26 May 2011

Correspondence to: Sten Hammarström, Department of Clinical Microbiology, Immunology, Umeå University, SE-90185 Umeå, Sweden, Tel: +46-(0)90-785-2345, Fax: +46-(0)90-785-2250, E-mail: sten.hammarstrom@climi.umu.se

Surgery is the primary treatment for CRC. Mesenteric lymph node metastasis is the single most important prognostic characteristic in CRC resected for cure, representing evidence of tumor cell dissemination beyond the primary location.[2–4] Approximately 50% of patients with lymph node involvement, Stage III (anyTN1-2M0) CRC, will experience disease recurrence.[2,4–7] Because a fraction (~25%) of patients without lymph node involvement, Stage I (T1-2N0M0) and Stage II (T3-4N0M0), also will develop recurrent disease after surgery, it is assumed that many of these patients have tumor cells in their lymph nodes not detected by conventional histopathological hematoxylin and eosin (H&E) staining.[2,4] Understaging may have several explanations, including insufficient number of retrieved lymph nodes for examination and inadequate sensitivity of the detection method. Of the different techniques investigated to improve the detection of disseminated tumor cells in lymph nodes, mRNA analysis by real-time quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) analysis showed great promise for the biomarkers carcinoembryonic antigen (CEA, also known as CEACAM5) and cytokeratin 20 (CK20).[8,9] When used with a specific probe and an RNA copy standard, this quantitative method is highly sensitive, specific and objective.

Accurate staging of CRC patients is also of value as a guide for selection of patients who may benefit from adjuvant

chemotherapy. Such treatment reduces the relative number of deaths by one-third in patients who undergo surgery for CRC in Stage III.[10,11] Adjuvant chemotherapy is, however, not regularly administered to Stage I and Stage II patients. Thus, it is important with improved selection criteria for identifying patients who may benefit from adjuvant chemotherapy and intense follow-up protocols.

Guanylyl cyclase C (GCC, also known as GUCY2C) and the stem cell marker CUB (complement protein subcomponents C1r/C1s, urchin embryonic growth factor, and bone morphogenic protein 1) containing protein 1 (CDCP1) are new biomarkers of potential interest for detecting disseminated tumor cells in CRC. Analysis of these markers may provide information complementary to that obtained by analysis of CEA or CK20. GCC is a cell surface associated glycoprotein that catalyzes the conversion of Guanosine-5′-triphosphate (GTP) to cyclic guanosine monophosphate (cGMP). Expression of GCC was reported to be restricted to the normal intestinal epithelium and colorectal tumor cells and to be a superior biomarker for detection of disseminated tumor cells in lymph nodes.[12–14] CDCP1 is a Type I transmembrane protein that contains three CUB domains. It was reported to be upregulated in colon tumors and expressed in human lung tumors and erythoblastoid stem cells.[15,16]

Ten to 20% of all CRC tumors are mucinous, which involves copious production of intestinal mucins, mainly mucin 2 (MUC2).[17,18] MUC2 is one of 17 mucins and the major mucin in the colon mucous layer covering the epithelium.[19] Mucinous tumors are defined as having more than 50% mucinous elements by histopathological examination.[17]

In our study, we have compared mRNA analysis of five biomarkers, CEA, CK20, GCC, CDCP1 and MUC2, for utility in detecting disseminated tumor cells in mesenteric lymph nodes staging and prediction of survival. The utility and predictive value of mRNA analysis was compared with conventional H&E histopathology analysis.

Material and Methods

Patients

Surgery for CRC was carried out in 174 patients [87 men and 87 women, median age 72 (range 51–90) years]. Sixteen tumors were located in the rectum and 158 in the colon. Seven of the 16 patients with rectal cancer received 25 Gy preoperative radiotherapy. A local radical resection of the tumors with wide lymph node dissection was carried out in all patients. According to the Tumor, Node, Metastasis (TNM) classification, 30 patients were in Stage I (T1-2N0M0), 79 in Stage II (T3-4N0M0), 47 in Stage III (anyTN1-2M0) and 18 had distant metastases, Stage IV (anyTanyNM1). Thirty-seven patients (4 in Stage II, 20 in Stage III and 13 in Stage IV) received postoperative chemotherapy. The tumor differentiation grade was poor, moderate and high in 10, 153 and 11 tumors, respectively. The most recent follow-up was performed January 1, 2011. No patient was lost for follow-up. The median follow-up time after surgery was 59 (range 17–131) months. Controls included 18 men and 6 women [median age 25 (range 10–61) years] undergoing colorectal surgery for ulcerative colitis (UC; $n = 18$), Crohn's colitis ($n = 4$), rectal prolapse ($n = 1$) and lipoma ($n = 1$).

Informed consent was obtained from the patients and in one case his parents. The Research Ethics Committee of the Medical Faculty, Umeå University, Sweden, approved the study.

Lymph nodes

Lymph nodes for mRNA analysis were dissected immediately from the surgically removed specimens and bisected under sterile conditions with separate knives to prevent RNA cross contamination. One half of each node was fixed in 10% buffered formalin for routine H&E staining and the other half was snap frozen in liquid nitrogen and stored at −70°C until RNA extraction. Lymph nodes [1–15 (median 2)] were obtained from each patient. Lymph nodes (517) were collected for biomarker mRNA and H&E analysis from the CRC patients (91 nodes from Stage I patients, 261 nodes from Stage II patients, 115 nodes from Stage III patients and 50 nodes from Stage IV patients). H&E staining only was performed on an additional 1,904 lymph nodes from CRC patients, giving a median of 13 nodes per patient (range 1–51 nodes/patient) that were analyzed by H&E. Lymph nodes (118) were collected from the control patients, 85 nodes from UC patients, 16 nodes from Crohn's colitis patients, 13 nodes from the patient with lipoma and 4 nodes from the patient with rectal prolapse.

Primary CRC tissue

In all, 87 samples from 57 tumors were analyzed for biomarker mRNA levels (17 samples were from tumors of 11 Stage I patients, 32 samples from tumors of 23 Stage II patients, 32 samples from tumors of 17 Stage III patients and 6 samples from tumors of 6 Stage IV patients). Eleven of these tumors were pT2, 39 tumors were pT3 and 7 tumors pT4. Tumor tissue samples, ∼0.5 × 0.5 × 0.5 cm in size, were collected from the primary tumor specimens immediately after resection, snap-frozen and kept at −70°C until RNA extraction. One to four samples were collected from each primary tumor.

Epithelial cells from colon tissue

Colonic epithelial cells (ECs) were isolated from apparently normal colons at the proximal or distal resection margins during surgical removal of tumors in CRC patients. The isolation procedure yields one fraction enriched in crypt ECs (crypt-ECs) and one fraction enriched in luminal ECs (luminal-ECs).[20]

Cell lines and peripheral blood mononuclear cells

Total RNA from the following human cell lines were analyzed: LS174T, HT29, T84, HCT8 and Caco2 (all colon carcinomas), Jurkat and Molt-4 (T cell lymphomas), B6 and KR4 (EBV-transformed B cell lines; a mixture of equal amounts of RNA from the two lines was used in the analyses), U266 (plasmacytoma), U937 (monocyte-like cell line), K562

(erythroblastoid cell line), HL60 (granulocyte cell line) and FSU (fibroblast cell line). Peripheral blood mononuclear cells (PBMCs) were isolated from peripheral blood of healthy adults by Ficoll-isopaque gradient centrifugation. PBMCs were activated *in vitro* by incubation with the anti-CD3 monoclonal antibody OKT3 (50 ng/mL) in HEPES-buffered RPMI 1640 supplemented with 0.4% human serum albumin. PBMCs from seven individuals were incubated with the stimulus in parallel cultures for 4, 7, 20, 48 and 72 hr, washed, pooled and RNA extracted.

RNA extraction

Total RNA was extracted from normal and tumor colon tissue, ECs, PBMCs and cell lines using the acid guanidine phenol chloroform method[21] by adding 0.5 mL of a solution containing 4 M guanidinium thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% sarcosyl and 0.1 M 2-mercaptoethanol per 25 mg tissue and up to 2.5×10^6 cells in the first homogenization step. Extracted RNA was dissolved in RNAse-free water containing the RNAse inhibitor RNAsin (1 U/μL; Promega, Madison, WI).

Real-time qRT-PCR

Real time qRT-PCR assays for quantitative determination of CEA, CK20, MUC2, GCC and CDCP1 mRNAs were developed in the laboratory using the Taqman EZ technology in which the 3′-primer is used for specific reverse transcription in each reaction (Taqman EZ RT-PCR; Applied Biosystems, Foster City, CA). This technique allows for detection of low amounts of specific mRNAs.[8] To securely exclude unspecific signals from genomic DNA, specific primer pairs were hybridized to different exons, and a dye-labeled probe was hybridized over the boundary between the two exons in the amplicon. Assays for CEA, CK20 and MUC2 mRNAs have been described.[8,9] Sequences for primers and probe for GCC mRNA were: forward primer 5′-CAAGCTGAAGGGTGACCGA-3′, reverse primer 5′-CAGGAGACAGCGTCAGAACAAG-3′ and probe 5′-CTTTTCAATGACCAGTACTT-3′, and for CDCP1 mRNA: forward primer 5′-CGGCTTCAGCATTGCAA-3′, reverse primer 5′-CCTTCACCCTCAAACACAGACT-3′ and probe 5′-CGCTCATCTATAAAACGTCTGTGCATCATC-3′. The reporter dye at the 5′-end of both probes was FAM. The quencher dye at the 3′-end was MGB for GCC and TAMRA for CDCP1. Emission from released reporter dye was measured by ABI Prism 7700 Sequence Detection System (Perkin-Elmer, Wellesley, MA). The RT-PCR profile for the GCC and CDCP1 assays was: 49°C for 2 min, 59°C for 30 min, 94°C for 5 min followed by 45 cycles of 93°C for 20 sec and 61°C for 1 min. RNA copy standards were prepared as described.[20] Serial dilutions of the respective RNA copy standard at concentrations from 10^2 to 10^7 copies/μL were included in each analysis and used as external standards for determination of concentrations in unknown samples. Analyses of unknown samples were carried out in triplicate and expressed as copies of mRNA/μL. The reproducibility of the qRT-PCR assays for the five biomarker-mRNAs was determined by performing six independent experiments in which 12 lymph nodes were analyzed for content of each biomarker mRNA. All lymph nodes were analyzed in triplicate with new dilutions of the RNA copy standard in each experiment. The coefficient of variation (C.V.) for the qRT-PCR method in all biomarker assays based on threshold PCR-cycles (ct-values) was <3.9% (range: 0.3–3.8%). The mean C.V.s for estimation of mRNA content covering 10^3 to 10^8 RNA copies per reaction were 25% for CEA, 9% for CK20, 11% for GCC, 22% for CDCP1 and 16% for MUC2. The concentration of 18S rRNA was also determined in each sample by real-time qRT-PCR according to the manufacturer's protocol (Applied Biosystems). 18S rRNA has been proven to be a stable and reliable housekeeping gene RNA suitable for normalizing mRNA levels among immune cells.[22] As no copy number standard is available for the 18S rRNA assay, 18S rRNA content was expressed as arbitrary units defined as the amount of 18S rRNA in 1 pg of total RNA extracted from polyclonally stimulated PBMCs.[22] Results are expressed as mRNA copies per unit of 18S rRNA, thus yielding directly comparable levels between different biomarkers.

Statistical analysis

Correlation between biomarker mRNA levels in lymph nodes was analyzed using the nonparametric Spearman correlation coefficient. Differences in expression levels and ratios between two patient groups were analyzed by two-tailed Mann-Whitney *U* test. The cut-off level for each biomarker was determined by dividing the clinical material into five groups of equal numbers of patients according to the biomarker level in their highest lymph node. The groups were compared with respect to disease-free survival using Cox regression analysis. From this analysis, the cut-off level was defined as the 80th percentile because, for all markers, the groups below the 80th percentile did not differ significantly in survival. Patients who died from causes other than CRC were considered as disease-free. Hazard ratios were determined using univariate as well as multivariate Cox regression analysis. Cumulative survival and mean survival time were determined using the Kaplan-Meier survival model in combination with the log rank test. Correlations between mRNA levels, differences in mRNA levels, differences in survival time and hazard ratios with a p value < 0.05 were considered to be statistically significant. The software used was SPSS version 18 (Chicago, IL).

Results

High and homogenous CEA mRNA expression in primary CRC tumors

The mRNA levels of the five biomarkers in 57 primary CRC tumors and in normal colon tissue are shown in Table 1. All tumors had detectable levels of mRNA for the five biomarkers. CEA was expressed at ~10- to 20-fold higher levels than CK20, GCC and CDCP1. MUC2 was expressed at the lowest level. Ten primary tumors were analyzed for homogeneity of biomarker mRNA levels. Tissue samples obtained

Table 1. Expression levels of CEA, CK20, GCC, CDCP1 and MUC2 mRNAs in primary CRC tumors, normal colon epithelium, CRC-, immune cell- and fibroblast cell lines

			Biomarker mRNA				
Source		*n*	CEA	CK20	GCC	CDCP1	MUC2
Primary CRC tumors		57	176 (112–281)[1]	18 (9–31)	7 (5–11)	7 (6–10)	1 (0.3–4)
Normal colon tissue		5	222 (133–307)	267 (188–377)	17 (14–32)	35 (15–46)	9 (6–25)
Normal colon crypt-ECs		5	261 (225–537)	295 (178–675)	6 (5–7)	4 (3–6)	32 (25–56)
Normal colon luminal-ECs		5	393 (212–769)	162 (104–377)	6 (3–7)	4 (2–6)	33 (17–77)
CRC cell lines	LS174T		328[2]	0.02	3.8	26.5	4.3
	HT29		32	85	0.004	14.9	0.01
	T84		33	33	12	3.2	0.5
	HCT8		32	0.05	0.05	11.6	0.02
	Caco2		3	7	5	23	0.04
PBMCs			0[3]	0	0.002	0.03	0
Activated PBMCs			0	0	0.003	1.6	0
T cell line	Jurkat		0	0	0.002	0.01	0
T cell line	Molt4		0	0	0.002	0.007	0
B cell lines	B6 + KR4		0	0	0.04	0.1	0
Plasma cell line	U266		0	0	0.02	0.004	0
Monocyte cell line	U937		0.005	0.003	0	0.03	0
Granulocyte cell line	HL60		0	0	0.00006	0.0004	0
Pre-erythrocyte cell line	K562		0	0.0003	0.001	11.1	0
Fibroblast cell line	FSU		0.0002	0.02	0.003	3.8	0.004

[1]mRNA copies/18S rRNA unit. Median and interquartile range from 25th to 75th percentile (IQR). [2]mRNA copies/18S rRNA unit. Mean of three determinations. [3]0, < 0.00001 mRNA copies/18S rRNA unit. Abbreviations: EC, epithelial cell; PBMC, peripheral blood mononuclear cells.

from four different sites within the same tumor revealed that CEA, CK20, GCC and CDCP1 were homogeneously expressed. MUC2, in contrast, showed intra-tumor variation (Supporting Information Fig. 1). There was no correlation between mRNA levels and tumor (T)-stage for any of the biomarkers (data not shown). Because of few highly differentiated tumors, it was not possible to evaluate whether biomarker levels were correlated to the degree of tumor differentiation. The biomarker mRNA levels in five established human colon carcinoma cell lines varied within one order of magnitude for CDCP1, two orders of magnitude for CEA and MUC2 as well as three to four for CK20 and GCC (Table 1). CEA, GCC and CDCP1 were expressed in normal colon tissue and in isolated normal colonocytes from the crypt and the free luminal surface at very similar levels to those of the primary CRC tumors, whereas both CK20 and MUC2 were present at about tenfold lower levels in primary tumor compared to normal colon tissue (Table 1).

No expression of CEA and MUC2 mRNAs in immune cells and fibroblasts

Because the biomarker analysis aims at identifying tumor cells in lymph nodes, it was essential to determine whether immune cells and fibroblasts express these markers. CEA and MUC2 were essentially not expressed in these cell types. CK20 was weakly expressed in fibroblasts and GCC in B cells and plasma cells. CDCP1 showed little specificity for tumor cells because the pre-erythrocyte cell line K562, fibroblasts and activated PBMCs expressed high levels of CDCP1 (Table 1). We calculated a specificity index, defined as the median value in primary CRC tumors/highest value of any immune cell or fibroblast. The index was 35,200 for CEA, 900 for CK20, 250 for MUC2, 175 for GCC and 0.6 for CDCP1.

Low levels of mRNA for CEA, CK20, GCC and MUC2 in lymph nodes from patients with benign intestinal disease

The mRNA levels for CEA, CK20, GCC and MUC2 were low but clearly detectable in lymph nodes of controls, generally <0.01 copies/18S rRNA unit, whereas the CDCP1 levels were two orders of magnitude higher (Fig. 1). CEA and CK20 levels in individual control nodes were strongly correlated to each other ($r = 0.49$, $p < 0.0001$) indicating that the same type of cells, exosomes and/or apoptotic bodies carry the mRNA detected in the two assays. Other combinations of markers showed no correlation or very weak correlation (data not shown).

Definition of clinically relevant cut-off levels for biomarkers by Cox regression model

Figure 1 shows the biomarker mRNA levels in lymph nodes of CRC patients where each patient is represented by the

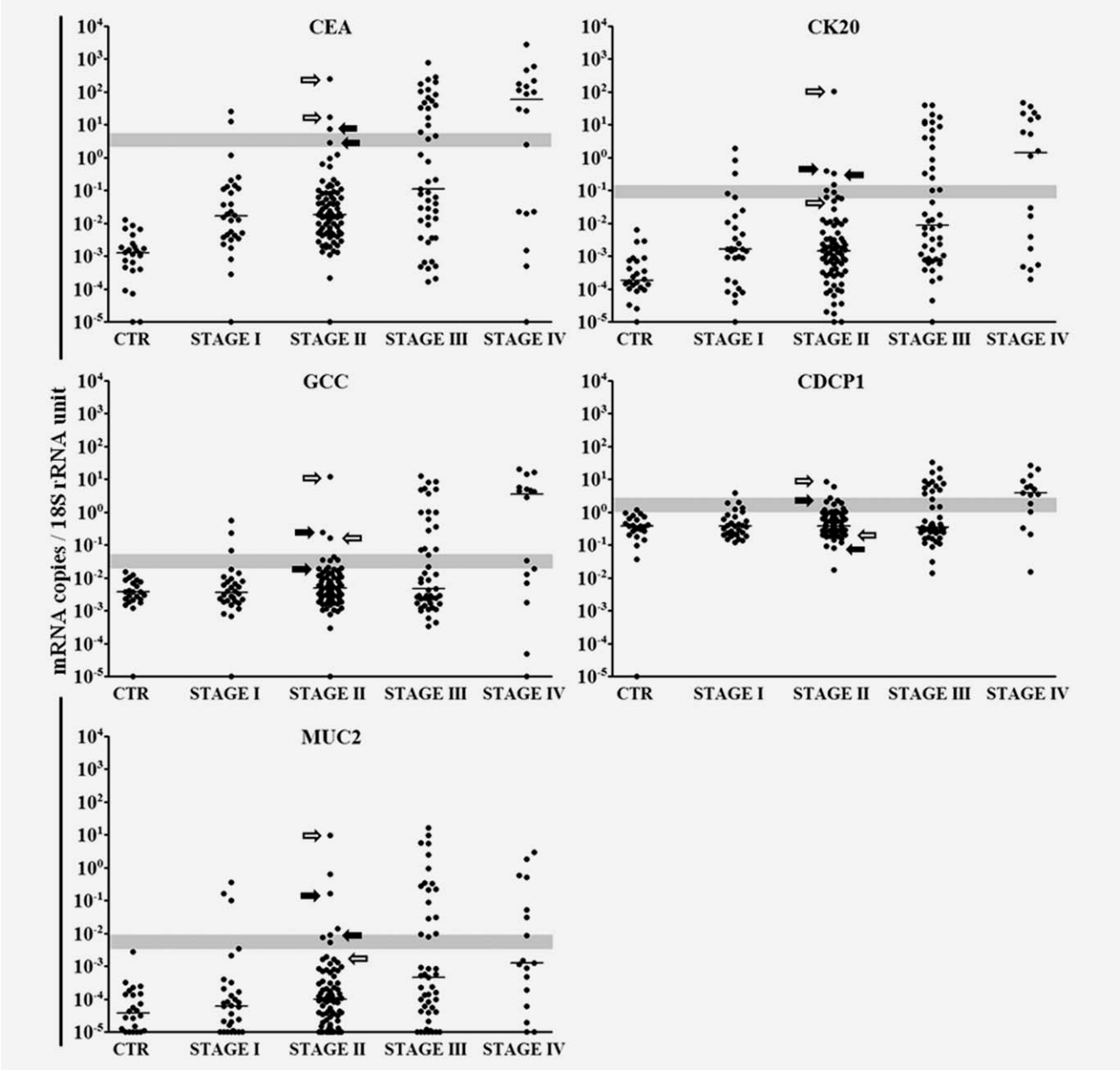

Figure 1. Biomarker mRNA levels in lymph nodes of Stages I–IV CRC patients and control patients (CTR). Each of the 174 CRC patients and 24 control patients is represented by the lymph node with the highest value of the respective mRNA species. The horizontal bars indicate median values for each group. The gray zones indicate the defined cut-off levels with experimental error.[9] Black arrows indicate CEA(+) patients who have died from recurrent CRC. Open arrows indicate CEA(+) patients who have died of unrelated disease.

node with the highest value. Patients are grouped according to TNM Stages I–IV. As expected, the median levels increased with stage. The increase was most pronounced for CEA and CK20 and only marginal for CDCP1. Most CDCP1 values were actually in the range of control nodes.

To define a clinically relevant cut-off level, *i.e.,* the mRNA level that best discriminates between the patients who get recurrent CRC and those who do not, for the different biomarker mRNAs, we made use of patients' survival data by using Cox regression analysis. During the follow-up time, 56 patients (5 Stage I, 11 Stage II, 24 Stage III and 16 Stage IV) had either died from CRC or were living with recurrent disease, whereas 118 patients were apparently cured or had died of causes unrelated to CRC. We defined the biomarker value at the 80th percentile as the cut-off level. The actual values are given in Table 2, which also shows the hazard ratios for recurrence comparing the marker positive [marker(+)] and marker negative [marker(−)] patient populations for each

Table 2. Comparative analyses of risk for recurrence of disease in biomarker positive, surgically treated CRC patients as determined by COX proportional hazard model and average survival time after surgery in biomarker(−)/(+) CRC patients as determined by cumulative survival according to Kaplan-Meier

Marker	Average no. of nodes analyzed per patient	Clinical cut-off	Marker(+) patients with recurrence of CRC		Marker(+) patients with no detected recurrence of CRC		Multivariate		Average survival time after surgery (months)			
			n	%	*n*	%	Hazard ratio (95% CI)	*p* value	Marker(−) patients	Marker(+) patients	Difference	*p* value
CEA	2	3.67[1]	24/56[2]	43	11/118[3]	9	4.60 (2.68–7.91)	<0.0001	99[4]	39	60	<0.0001
CK20	2	0.099	23/56	41	13/118	11	3.87 (2.25–6.66)	<0.0001	98	42	56	<0.0001
GCC	2	0.035	20/54	37	13/117	11	3.58 (2.04–6.27)	<0.0001	97	43	54	<0.0001
CDCP1	2	1.83	17/53	32	16/116	14	2.61 (1.46–4.66)	0.001	94	62	32	0.001
MUC2	2	0.0062	20/55	36	14/117	12	3.46 (1.98–6.03)	<0.0001	96	52	44	<0.0001
H&E	2[5]	–	23/56	41	9/118	8	4.59 (2.67–7.90)	<0.0001	98	36	62	<0.0001
H&E	13[6]	–	37/56	66	23/118	20	6.35 (3.57–11.28)	<0.0001	107	40	67	<0.0001

[1]mRNA copies/18S rRNA unit. [2]Number of marker positive patients/total number of patients with recurrence of CRC. [3]Number of marker positive patients/total number of patients with no detected recurrence of CRC. [4]Average survival time after surgery as calculated by cumulative survival analysis according to Kaplan Meier. [5]The same lymph nodes as analyzed for biomarker mRNA. [6]The total number of lymph nodes analyzed by H&E.

biomarker. As can be seen, there is a 4.6-fold greater risk of recurrence for CEA(+) patients than for CEA(−) patients. The other biomarkers gave lower hazard ratios.

Patients (43% and 41%) who had died from CRC or had recurrent disease were CEA(+) and CK20(+), respectively, and thus should be considered correctly identified (Table 2). For GCC, MUC2 and CDCP1, the corresponding values were 37%, 36% and 32%, respectively. The percentage of marker(+) patients without recurrent disease, "false positives" was almost the same for all five biomarkers, varying from 9 to 14%, with CEA giving the lowest value (Table 2). Biomarker analysis identified totally 26 patients with recurrent CRC, of whom CEA identified 24. One of the CEA(−) patients with recurrent CRC was identified by CK20 [CEA(−)CK20(+)GCC(−)MUC2(−)] and the other one by MUC2 [CEA(−)CK20(−)GCC(−)MUC2(+)]. Both patients had CEA values just below the cut-off. To determine which of the biomarkers was best suited to predict tumor recurrence, we applied Cox multivariate analysis. CEA mRNA was statistically the strongest indicator ($p = 0.015$). Combined biomarker analysis was also investigated. Combination of CEA and CK20 gave the best results with increased sensitivity for identification of patients with recurrent CRC from 43 to 45% but at the same time the specificity decreased in that the percentage of marker(+) patients with no recurrent disease increased from 9 to 13% (15 patients of 118). Thus, over-all no improvement in predicting outcome was gained compared to analysis of CEA alone.

Interestingly, CEA mRNA analysis gave closely similar results as H&E staining of sections from the same lymph nodes (Table 2). On average, 13 lymph nodes/patient were analyzed in routine histopathology. As expected, the number of correctly identified patients with recurrent CRC increased when a larger number of H&E stained nodes were evaluated, *i.e.,* from 41 to 66%. Perhaps somewhat unexpected, however, was the finding that the percentage of "false positives" also increased, *i.e.,* from 8 to 20% (Table 2).

A relatively large fraction of Stage III patients had biomarker mRNA values below the cut-off level (Fig. 1). This is partly due to the fact that for only 21 of the 47 Stage III patients did we have access to H&E positive nodes for mRNA analysis. In the remaining 26, only H&E(−) nodes were available for mRNA analysis.

Best discrimination of survival time after surgery between biomarker(+) and biomarker(−) CRC patients is achieved by CEA mRNA

Figure 2 shows the cumulative survival according to Kaplan-Meier for the biomarker(+) and biomarker(−) patient groups using the defined cut-off levels for the five markers. For comparison, we have included the results for H&E(+) and H&E(−) patients using exactly the same material. Table 2 summarizes the mean survival times for the biomarker(+) and biomarker(−) populations for each marker. For the CEA(−) population, the survival time was 99 months

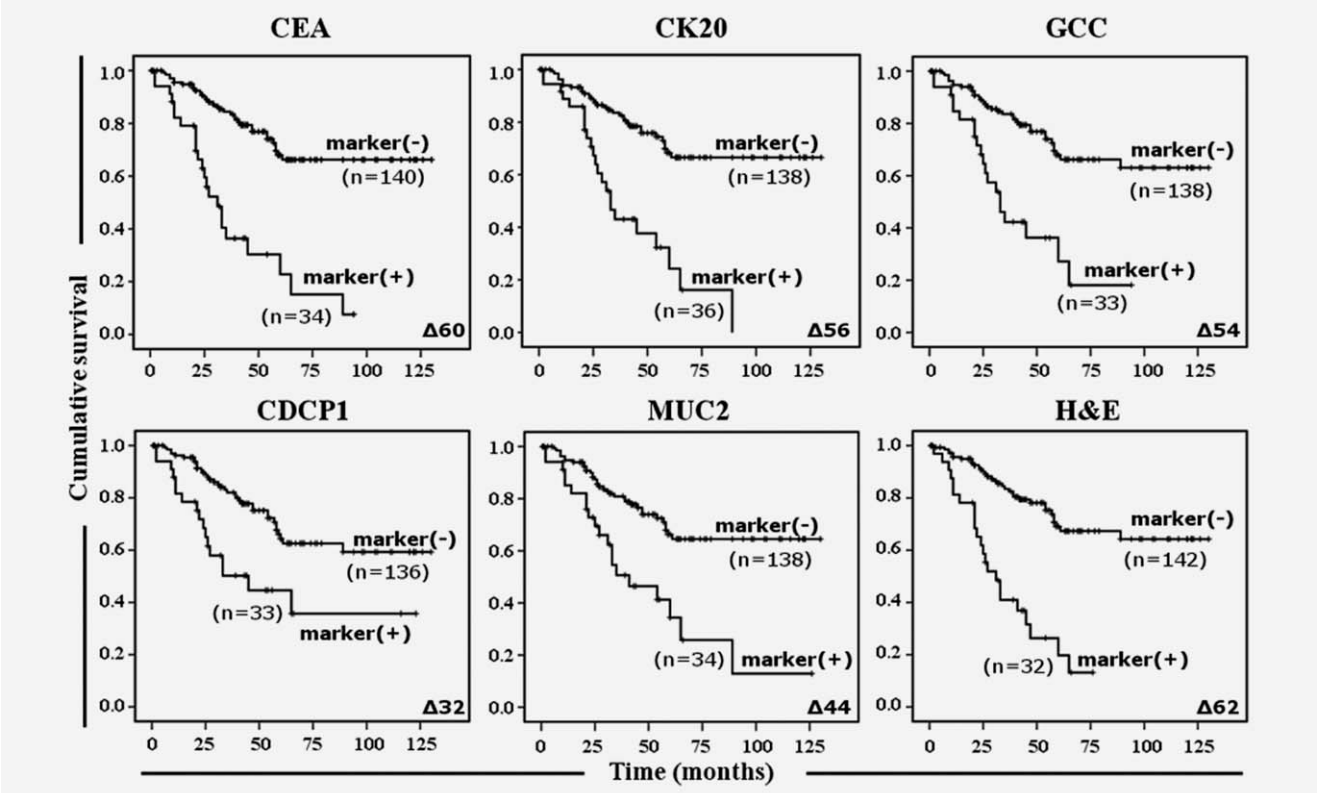

Figure 2. Cumulative survival curves according to Kaplan-Meier for lymph node biomarker(+) and lymph node biomarker(−) CRC patients using the clinically optimal cut-off levels for the five markers to discriminate between the two groups. For comparison, the corresponding curves based on H&E staining of the same nodes is included (on average two nodes per patient). The difference in average survival time between the marker(+) and marker(−) groups is given as a delta value (in months) for each marker. *n*, numbers of patients in the marker(+) and marker(−) groups, respectively.

compared to 39 months for the CEA(+) population. The difference of 60 months, *i.e.,* 5 years, was highly significant ($p < 0.001$). For the other biomarkers, the difference in mean survival time was shorter; although for CK20 and GCC, it was quite close, 56 and 54 months, respectively. Notably, CEA gave closely similar results as classification by H&E (Table 2).

We also investigated if use of multiple markers would improve the model's ability to predict survival between the marker(+) and marker(−) groups. However, in no case did any combination perform better than CEA alone (data not shown).

CEA mRNA has the best prognostic power for the detection of recurrence in CRC

The prognostic value of the five biomarkers was investigated by determining the fraction of marker(+) and marker(−) patients who had recurrence in CRC 1 year after surgery, 2 years after surgery, etc. (Fig. 3). Two groups of well-separated curves are seen. The percentage of recurrences increased with time in both groups and, for the marker(−) group, all five

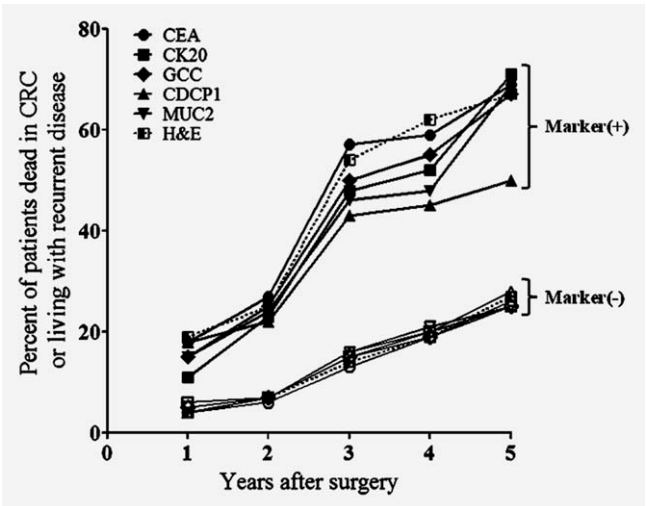

Figure 3. Proportion of CRC patients with recurrent disease in marker positive [marker(+)] and marker negative [marker(−)] groups as a function of time after surgery. (●), CEA mRNA; (■), CK20 mRNA; (◆), GCC mRNA; (▲), CDCP1 mRNA; (▼), MUC2 mRNA; (◧) and dotted line, H&E-staining.

biomarkers and H&E gave the same result (Fig. 3). In the marker(+) group, however, CEA and H&E identified a larger proportion of patients with CRC recurrence than the other biomarkers. The largest difference was seen 3 and 4 years after surgery.

Biomarker positivity identified Stage II patients with CRC recurrence

Six of 109 (5.5%) Stages I and II patients were CEA(+). The corresponding values for CK20, GCC and MUC2 were 8.2%, 9.2% and 9.2%, respectively. CDCP1 is not considered because of its poor specificity for colon epithelial/tumor cells. Interestingly, all four markers identified the same two patients who had died from recurrent CRC (Fig. 1). Of the remaining four CEA(+) patients, two have died from unrelated diseases and two Stage I patients were alive 24 and 43 months after surgery (Fig. 1).

A high MUC2/CEA mRNA ratio in lymph nodes of CRC patients indicates a better prognosis

Byrd and Bresalier reported that patients with mucinous tumors tend to have a better prognosis than patients with nonmucinous tumors.[17] Because MUC2 is the major mucin in the colon, we investigated whether the MUC2/CEA mRNA ratio in lymph nodes could identify patients who died from CRC or were living with recurrent disease during the follow-up period (Group A). The group of patients who were living without recurrent CRC (Group B) had significantly higher MUC2/CEA ratio than those with recurrent disease ($p < 0.05$; Fig. 4*a*). This was even more pronounced if only MUC2(+) patients were considered ($p < 0.01$; Fig. 4*b*). In the primary tumor, this difference was not observed (Fig. 4*c*). By dividing the MUC2(+) patients into three groups using the median MUC2/CEA ratios of Groups A and B (3.3×10^{-3} and 59.5×10^{-3}, respectively) as cut-offs, we could establish high-, intermediate- and low-risk groups. The high-risk group of patients was characterized by a low MUC2/CEA ratio ($<3.3 \times 10^{-3}$) and a significantly shorter average survival time compared to the intermediate- and low- risk groups, *i.e.,* 22, 46 and 61 months, respectively ($p = 0.011$; Fig. 4*d*). One of the marker(+) Stage II patient with recurrent CRC belonged to the high-risk group (Figs. 1 and 4*d*).

Because the CEA and CK20 levels in lymph nodes of CRC patients were correlated, we also calculated the MUC2/CK20 ratios for the different groups. The same trend was seen, but the result was weaker and reached statistical significance ($p < 0.05$) only for the MUC2(+) patients. Thus, the MUC2/CEA ratio is the most promising predictor of survival after surgery.

Discussion

The primary goal of our study was to determine, which of five promising biomarker mRNAs[9,12–14,23,24] for regional lymph node analysis was the best in identifying CRC patients with poor prognosis after surgery. To this end, we defined a clinically relevant cut-off level for each marker using Cox regression model on patients survival data. Notably, for CEA and CK20, this level was clearly higher than the highest level of lymph nodes from control patients. CEA gave the highest hazard ratio (4.60) for risk of recurrence and the largest difference in average survival time for marker(+) compared with marker(−) patients (60 months). Second best was CK20 with a hazard ratio of 3.87 and a difference in average survival time of 56 months. GCC came in third. We found that the sensitivity increased but the specificity decreased in predicting outcome when CK20 and MUC2 analysis was added in combination with CEA analysis. Moreover, CEA showed the highest tissue specificity (specificity index 35,200), the highest expression level per colon cancer cell and the smallest variation in expression level between and within primary CRC tumors. Thus, we conclude that CEA mRNA is the marker of choice if used as a single marker.

First, CEA mRNA analysis of lymph nodes gave the most similar result to that obtained by histopathology when the same lymph nodes were compared. Extended comparative studies are needed to determine whether CEA mRNA assays could eventually replace histopathology to determine lymph node status in TNM staging. Strengths of CEA mRNA analysis are foremost the ability to combine it with mRNA assays for other biomarkers, *e.g.,* MUC2 and still unidentified progression markers, and that the procedure is objective, quantitative and amendable for automation. Furthermore, a larger volume of each lymph node is analyzed, thereby significantly increasing the sensitivity for detection of disseminated tumor cells. Most likely biomarker mRNA analysis will, when automated, be a more cost-effective method to determine regional lymph node status than histopathology. A drawback is that it is difficult in the present clinical routine to collect a large number of lymph nodes from fresh specimens being representative for the regional lymphatic field.

Second, four of the five biomarkers (CEA, CK20, GCC and MUC2) were able to identify the two Stage II patients who succumbed from CRC within the observation period. Here, we can safely conclude that these patients were missed by routine histopathology. Possibly, these patients would have benefited from adjuvant chemotherapy. Whether lack of detection by routine histopathology was due to the small volume analyzed by H&E as compared to biomarker mRNA (a few percent *vs.* 50%),[24] the difficulty in identifying single or small groups of tumor cells by H&E, other factors or a combination remain to be determined. It is well established that regional lymph nodes from a significant proportion of Stage I and Stage II patients contain micrometastases or isolated tumor cells as detected by anti-CEA or anti-CK20 antibodies in immunohistochemistry (unpublished results).[8,23,25,26] Two Stage II patients with high CEA, CK20 and GCC levels (open arrows in Fig. 1) died from diseases unrelated to CRC. Possibly, their CRC may also have recurred. The proportion of biomarker positive Stage I + Stage II patients varied from 5.6 (CEA) to 9.2%.

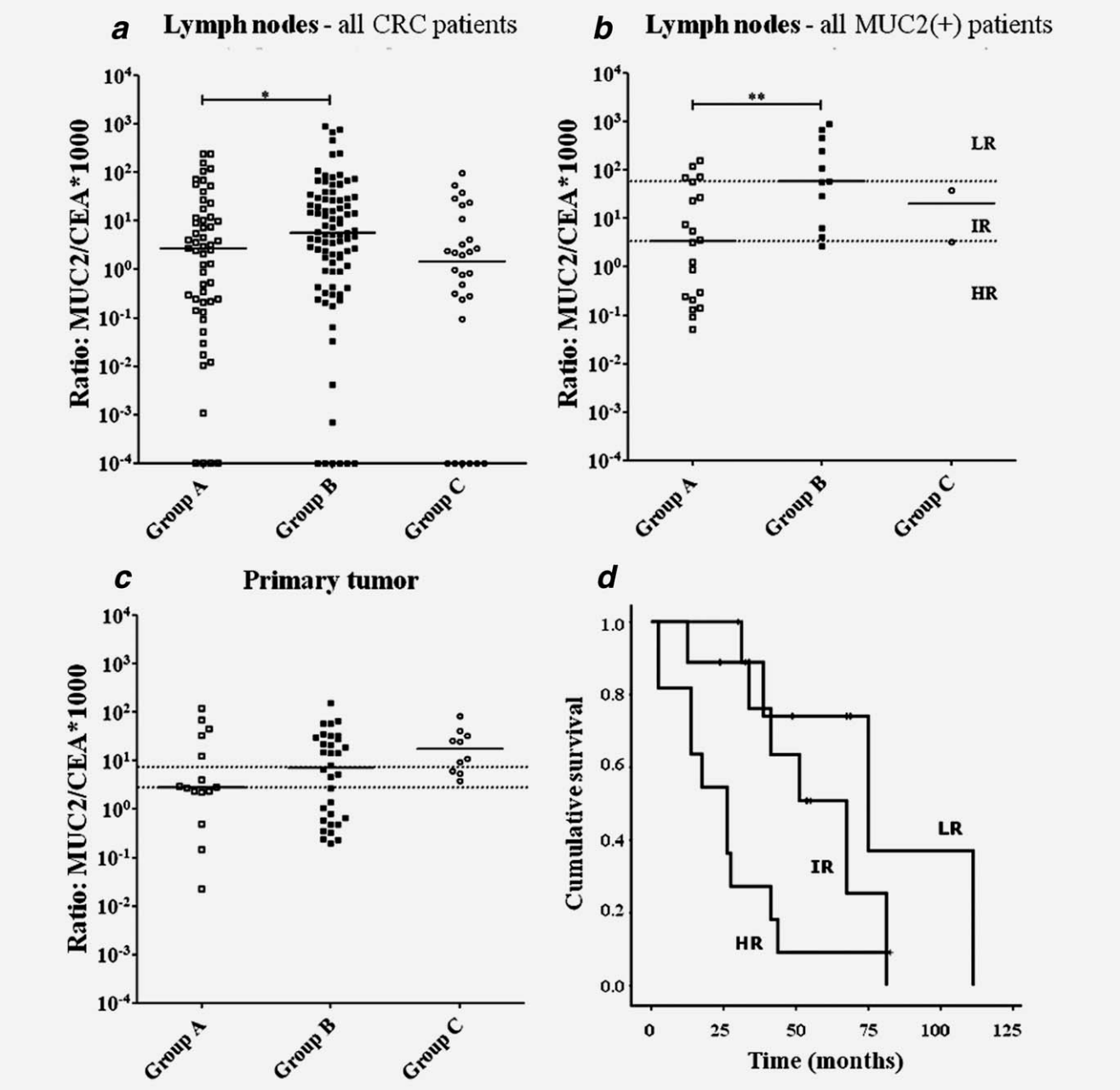

Figure 4. MUC2/CEA mRNA ratio in lymph nodes and primary tumors from CRC patients who have died from CRC or had recurrence within the follow-up time after surgery (Group A), CRC patients who are living without recurrent disease at the end of the observation period (Group B) and CRC patients who have died from other causes than CRC (Group C). (*a*) All 174 CRC patients, Stages I–IV. (*b*) All 35 MUC2(+) CRC patients, Stages I–IV. Dotted lines indicate the medians for Groups A and B, dividing the patients into low (LR), intermediate (IR) and high-risk (HR) groups. (*c*) Fifty-seven primary CRC tumors from Stages I–IV patients. (*d*) Cumulative survival curves according to Kaplan-Meier for the LR, IR and HR groups of patients defined in (*b*). Median ratios are indicated as horizontal bars in (*a–c*). **p* value $<$ 0.05; ***p* value $<$ 0.01.

Third, we found that CRC patients with MUC2(+) lymph nodes had a poor prognosis if their MUC2/CEA mRNA ratio was low. This probably reflects that tumors with a low MUC2/CEA ratio are poorly differentiated. Such tumors generally have a poorer prognosis than well-differentiated tumors. Determination of the MUC2/CEA mRNA ratio may therefore become useful in clinical practice for selecting patients for adjuvant therapy. In this case, it would be advisable to only consider MUC2(+) patients, thus excluding patients with lymph nodes that essentially lack tumor cells. It

is worth mentioning that these high-risk CRC patients could not be identified by analysis of the primary tumor, suggesting that it is the properties of the disseminating tumor cells that are of importance.

Fourth, neither biomarker analysis nor histopathology was able to detect all patients who succumbed to CRC during the observation time after surgery. Based on an average of 2 lymph nodes per patient, CEA analysis and H&E-staining detected about 40% of these patients. It is reasonable to assume that the proportion of "correctly" identified patients would have increased if a larger number of lymph nodes had been subjected to biomarker analysis. Indeed, by H&E staining, the proportion of correctly identified patients increased to 66% if the number of lymph nodes analyzed per patient was increased. However, it will probably not be possible to achieve 100% detection because some tumors may have spread by the hematogeneous route. At first glance, it may seem like a large margin of error that 10% of the marker(+) patients did not develop recurrent CRC within the observation period. A closer scrutiny of these patients, which essentially were the same individuals for all markers, revealed that three of them had died of causes unrelated to CRC and three may go on to develop recurrent disease, as they were evaluated only 21–26 months after surgery. Finally, three patients living 54, 54 and 39 months after surgery were all treated with adjuvant chemotherapy and potentially cured.

In our study, we show that CEA mRNA analysis of lymph nodes collected at surgery from patients with CRC can successfully be used to ascertain the regional lymph node status with at least the same degree of precision as with H&E-staining. An advantage of CEA mRNA detection over histopathology was the identification of a group of Stage II CRC patients with tumor cells in their lymph nodes. Moreover, additional prognostic information was obtained if the MUC2/CEA mRNA ratio was determined—a low ratio indicating poor prognosis.

References

1. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, Radich J, Anderson G, Hartwell L. The case for early detection. *Nat Rev Cancer* 2003;3:243–52.
2. Iddings D, Ahmad A, Elashoff D, Bilchik A. The prognostic effect of micrometastases in previously staged lymph node negative (N0) colorectal carcinoma: a meta-analysis. *Ann Surg Oncol* 2006;13:1386–92.
3. Chang GJ, Rodriguez-Bigas MA, Skibber JM, Moyer VA. Lymph node evaluation and survival after curative resection of colon cancer: systematic review. *J Natl Cancer Inst* 2007;99:433–41.
4. Nicastri DG, Doucette JT, Godfrey TE, Hughes SJ. Is occult lymph node disease in colorectal cancer patients clinically significant? A review of the relevant literature. *J Mol Diagn* 2007;9:563–71.
5. Hashiguchi Y, Hase K, Ueno H, Mochizuki H, Kajiwara Y, Ichikura T, Yamamoto J. Prognostic significance of the number of lymph nodes examined in colon cancer surgery. *Ann Surg* 2010;251:872–81.
6. Swanson RS, Compton CC, Stewart AK, Bland KI. The prognosis of T3N0 colon cancer is dependent on the number of lymph nodes examined. *Ann Surg Oncol* 2003;10:65–71.
7. Bilchik AJ, Hoon DS, Saha S, Turner RR, Wiese D, DiNome M, Koyanagi K, McCarter M, Shen P, Iddings D, Chen SL, Gonzalez M, et al. Prognostic impact of micrometastases in colon cancer: interim result of a prospective multicenter trial. *Ann Surg* 2007;246:568–75.
8. Öberg Å, Lindmark G, Israelsson A, Hammarström S, Hammarström M-L. Detection of occult tumor cells in lymph nodes of colorectal cancer patients using real-time quantitative RT-PCR for CEA and CK20 mRNAs. *Int J Cancer* 2004;111:101–10.
9. Ohlsson L, Hammarström M-L, Israelsson A, Näslund L, Öberg Å, Lindmark G, Hammarström S. Biomarker selection for detection of occult tumor cells in lymph nodes of colorectal cancer patients using real-time quantitative RT-PCR. *Br J Cancer* 2006;95:218–25.
10. Moertel CG, Fleming TR, MacDonald JS, Haller DG, Laurie JA, Tangen CM, Ungerleider JS, Emerson WA, Tormey DC, Glick JH, Veeder MH, Mailliard JA. Fluorouracil plus levamisole as effective adjuvant therapy after resection of stage III colon carcinoma: a final report. *Ann Intern Med* 1995;122:321–6.
11. Edge SB, Byrd DR, Compton CC, Fitz AG, Greene FL, Trotti A, eds. AJCC cancer staging manual, 7th edn. Philadelphia: Springer, 2010.
12. Frick GS, Pitari GM, Weinberg DS, Hyslop T, Schulz S, Waldman SA. Guanylyl cyclase C: a molecular marker for staging and postoperative surveillance of patients with colorectal cancer. *Expert Rev Mol Diagn* 2005;5:701–13.
13. Schulz S, Hyslop T, Haaf J, Bonaccorso C, Nielsen K, Witek ME, Birbe R, Palazzo J, Weinberg D, Waldman SA. A validated quantitative assay to detect occult micrometastases by reverse transcriptase-polymerase chain reaction of guanylyl cyclase C in patients with colorectal cancer. *Clin Cancer Res* 2006;12:4545–52.
14. Waldman SA, Hyslop T, Schulz S, Barkun A, Nielsen K, Haaf J, Bonaccorso C, Li Y, Weinberg DS. Association of GUCY2C expression in lymph nodes with time to recurrence and disease-free survival in pN0 colorectal cancer. *J Am Med Assoc* 2009;301:745–52.
15. Büring HJ, Kuci S, Conze T, Rathke G, Bartolović K, Grünebach F, Scherl-Mostageer M, Brümmendorf TH, Schweifer N, Lammers R. CDCP1 identifies a broad spectrum of normal and malignant stem/progenitor cell subset of hematopoietic and nonhematopoietic origin. *Stem Cells* 2004;22:334–43.
16. Wortman A, He Y, Deryugina EI, Quigley JP, Hooper JD. The cell surface glycoprotein CDCP1 in cancer—insights, opportunities, and challenges. *IUBMB Life* 2009:61:723–30.
17. Byrd JC, Bresalier RS. Mucins and mucin binding proteins in colorectal cancer. *Cancer Metastasis Rev* 2004;23:77–99.
18. Park ET, Oh HK, Gum JR, Crawley SC, Kakar S, Engel J, Leow CC, Gao WQ, Kim YS. HATHI expression in mucinous cancers of the colorectum and related lesions. *Clin Cancer Res* 2006;12:5403–10.
19. Derrien M, van Passel MWJ, van de Bovenkamp JHB, Schipper RG, de Vos WM, Dekker J. Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes* 2010;1:254–68.
20. Fahlgren F, Hammarström S, Danielsson Å, Hammarström M-L. Increased expression of antimicrobial peptides and lysozyme in colonic epithelial cells of patients with ulcerative colitis. *Clin Exp Immunol* 2003;131:90–101.
21. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–9.
22. Bas A, Forsberg G, Hammarström S, Hammarström M-L. Utility of the

housekeeping genes 18S rRNA, β-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand J Immunol* 2004;59:566–73.

23. Nordgård O, Oltedal S, Kørner H, Aasprong OG, Tjensvoll K, Gilje B, Heikkilä R. Quantitative RT-PCR detection of tumor cells in sentinel lymph nodes isolated from colon cancer patients with an *ex vivo* approach. *Ann Surg* 2009;249:602–7.
24. Haince J-F, Houde M, Beaudry G, L'espérance S, Garon G, Desaulniers M, Hafer LJ, Heald JI, Lyle S, Grossman SR, Têtu B, Sargent DJ, et al. Comparison of histopathology and RT-qPCR amplification of guanylyl cyclase C for detection of colon cancer metastases in lymph nodes. *J Clin Pathol* 2010;63:530–7.
25. van der Zaag ES, Buskens CJ, Kooij N, Akol H, Peters HM, Bouma WH, Bemelman WA. Improving staging accuracy in colon and rectal cancer by sentinel lymph node mapping: a comparative study. *Eur J Surg Oncol* 2009;35:1065–70.
26. Davies M, Arumugam PJ, Shah VI, Watkins A, Roger Morgan A, Carr ND, Beynon J. The clinical significance of lymph node micrometastasis in stage I and II colorectal cancer. *Clin Transl Oncol* 2008;10:175–9.