CLINICAL GASTROENTEROLOGY AND HEPATOLOGY 2008;6:1122–1128

Improved Diagnosis of Colorectal Cancer Using a Combination of Fecal Occult Blood and Novel Fecal Protein Markers

JOHANN KARL,* NORBERT WILD,* MICHAEL TACKE,* HERBERT ANDRES,* URSULA GARCZAREK,‡ WOLFGANG ROLLINGER,* and WERNER ZOLG*

*Department of New Technologies, and the ‡Department of Biostatistics, Professional Diagnostics, Roche Diagnostics GmbH, Penzberg, Germany

Background & Aims: Annual testing for fecal occult blood is recommended as first-line screening for the detection of colorectal cancer (CRC), but is affected by limited sensitivity. We initiated a proteomics-based search for novel biomarkers to improve the sensitivity of detection of CRC in stool samples. Methods: Six markers, including immunologic fecal occult blood test (iFOBT), were evaluated in a collective of 551 samples (186 CRC, 113 advanced adenoma, and 252 control patients) to establish the diagnostic performance of each marker and marker combinations. Results: We tested the known stool markers hemoglobin (iFOBT), hemoglobin-haptoglobin, calprotectin, carcinoembryogenic antigen, and the novel fecal markers tissue inhibitor of metalloproteinase-1 (TIMP-1) and S100A12. The best diagnostic performance was found for S100A12 with an area under the curve of 0.95, followed by TIMP-1 (0.92), hemoglobin-haptoglobin (0.92), hemoglobin (0.91), calprotectin (0.90), and carcinoembryogenic antigen (0.66). By using Bayes logistic regression as a mathematic model, the highest sensitivity (88%) for the detection of CRC at 95% specificity was obtained with the marker pair S100A12 and hemoglobin-haptoglobin. Increasing the specificity to 98%, the combination of S100A12, hemoglobin-haptoglobin, and TIMP-1 resulted in a sensitivity of 82%, with the highest increase of sensitivity found in early tumor stages (international union against cancer stage I: 74% sensitivity vs 57% of the best single marker). Conclusions: Depending on the specificity selected, a marker pair, S100A12 and hemoglobin-haptoglobin, or a triple combination including TIMP-1, allowed the detection of CRC at significantly higher rates than can be obtained with iFOBT alone.

Colorectal cancer (CRC) is one of the most prevalent cancers worldwide and the lifetime risk is almost 6%.[1] Early detection is clearly a key factor in reducing mortality from CRC.[2] Several screening regimens for CRC are recommended, including colonoscopy, fecal occult blood testing (FOBT), and fecal DNA analysis. Although colonoscopy remains the gold standard for the detection of colon lesions, compliance is low owing to uncomfortable and unpleasant preparation procedures. Other limitations of colonoscopy for primary screening are the risk of complications, costs, and access. In contrast, stool-based testing is well accepted, despite limitations such as low sensitivity and dietary influences. The commonly used guaiac-based FOBT is an effective screening tool when used FOBT (iFOBT) assays, which are specific for human hemoglobin, and eliminate the need for dietary restrictions and have a similar or better sensitivity. A recent study including 21,805 asymptomatic Japanese patients reported a sensitivity for iFOBT of 65.8% for detecting cancer (specificity, 95%).[6] Notwithstanding improved sensitivity compared with guaiac-based FOBT, about one third of invasive cancer remained undetected in this study. New tools improving the sensitivity of CRC screening are therefore needed. Recently, a second-generation fecal DNA test with improved performance has been reported.[7] Fecal DNA testing and virtual colonoscopy now are included in the joint guidelines for CRC screening.[8]

Applying proteomics approaches to identify new screening markers we analyzed the protein expression in colon tissue and found strongly increased expression of S100A12 in CRC.[9] The objective of the present study was to examine the clinical performance of fecal S100A12 and other selected biomarkers for early detection of CRC in stool samples in comparison with iFOBT and to evaluate if marker combinations can improve the sensitivity further.

Materials and Methods

Study Design

Stool samples were collected prospectively in 2 European multicenter studies. The first study recruited patients at gastroenterology units in an average-risk screening population that underwent a preventive check by colonoscopy. Patients with symptoms of gastrointestinal events such as rectal bleeding, recent change in bowel habits, or lower abdominal pain, and if FOBT testing was performed before colonoscopy, were excluded. From each participant a colonoscopy was performed at the participating centers with preparation and sedation used at each site. The stool samples had to be collected before colonoscopy. In control patients collection also was permitted if performed more than 3 days after colonoscopy. The size and location of each lesion were recorded. A pathologist examined each surgical resection specimen on site to determine the diag-

Abbreviations used in this paper: BLR, Bayes Logistic Regression; CEA, carcinoembryogenic antigen; CRC, colorectal cancer; FOBT, fecal occult blood test; iFOBT, immunologic fecal occult blood test; TIMP-1, tissue inhibitor of metalloproteinase-1; UICC, international union against cancer.

Table 1. Basic Characteristics of Patient Collectives and Measured Marker Concentrations

	Patients, (n)	Mean age (±SD), (y)	Sex	
			Female	Male
Controls	252	63.0 ± 8.0	151	101
Healthy[a]	132	62.3 ± 6.8	81	51
Hemorrhoids	28	60.1 ± 7.1	13	15
Diverticulosis	73	64.7 ± 9.5	46	27
Hyperplastic polyps	14	67.9 ± 9.9	8	6
Other bowel diseases	5	59.2 ± 6.7	3	2
Advanced adenoma	113	66.8 ± 8.5	48	65
CRC collective I (all stages)[b]	101	68.4 ± 11.5	48	53
UICC 0/I	1/22	65.0 ± 10.0	8	15
UICC II	27	73.1 ± 10.9	14	13
UICC III	12	70.9 ± 12.3	8	4
UICC IV	23	69.6 ± 10.3	12	11
Without staging	16	61.9 ± 12.6	6	10
CRC collective II (all stages)[c]	85	64.0 ± 11.8	44	41

NOTE. Median values shown.
[a]No evidence of bowel disease.
[b]No CRC patient underwent FOBT or had visible blood in his/her stool before colonoscopy.
[c]CRC patients underwent a colonoscopy because of a positive FOBT or because of visible blood in their stool; in 11 patients the reason for colonoscopy was unknown.

nosis and the respective staging. Because of the low incidence of approximately 0.4% CRC patients within the preventive screening population, cancer patients were additionally recruited in a second prospective study at different surgery units without any restrictions regarding symptoms or FOBT testing. In this study the stool samples were collected before surgery. The diagnosis of CRC was confirmed by pathologic staging of each patient by pathologists on site. The research protocols for both studies were reviewed and approved by the appropriate ethics committees and all participants gave written informed consent.

Clinical samples from both multicenter studies were compiled for the evaluation. Group A comprised the control cohort with 252 patients from study I. All patients with adenoma or inflammatory bowel diseases were excluded. Group B comprised the advanced adenoma cohort containing 113 patients from study I and study II with any lesion containing high-grade dysplasia, villous or tubovillous architecture, or tubular adenoma with a diameter of at least 1 cm. Group C comprised the CRC cohort with 186 CRC patients from study I and study II (Table 1 and Figure 1). To avoid a positive bias for the iFOBT assay, the cancer patients were divided into collective I (no FOBT testing or visible blood in stool), and collective II (no restrictions applied). Only 4 CRC patients from collective I had an additional inflammatory bowel disease. To assess the influence of tumor localization on the diagnostic result, collective I was subdivided into right-sided CRC (cecum to colon transversum) and left-sided CRC (flexura lienalis to rectum).

Stool Sample Collection

Each participant provided 2 different portions of approximately 1 g of feces from one bowel movement using a stool collection tube (identification number 80.623.022; Sarstedt, Nümbrecht, Germany). Subjects were given detailed −20°C within 24 hours of collection and transferred within 2 weeks to a −70°C freezer. Stool samples were transported on dry ice to our laboratory and stored at −70°C.

Immunoassays

Six immunoassays were measured: Hemoglobin (RIDASCREEN Hemoglobin, R-Biopharm, Darmstadt, Germany), Hemoglobin-haptoglobin (RIDASCREEN Hemoglobin-Haptoglobin), calprotectin (calprotectin test; Nova Tec Immundiagnostica GmbH, Dietzenbach, Germany), tissue inhibitor of metalloproteinase-1 (TIMP-1) (Quantikine human TIMP-1; R&D Systems, Minneapolis, MN), and carcinoembryonic anti-

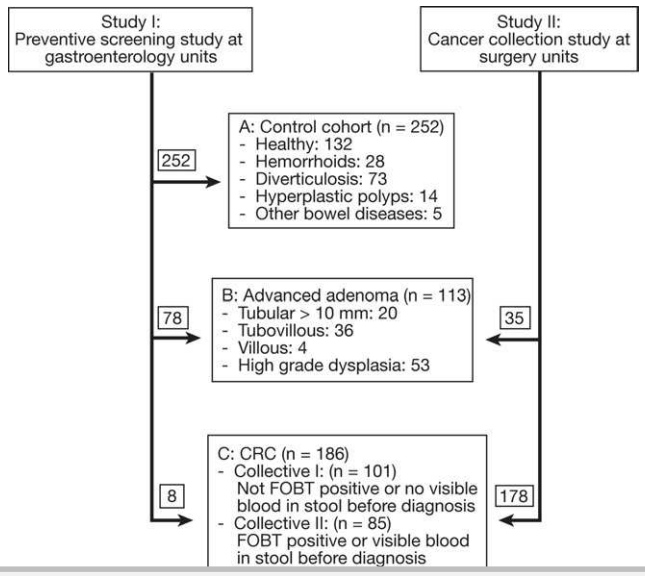

Table 1. Continued

S100A12 (*ng/g*)	Hemoglobin (*μg/g*)	Hemoglobin-haptoglobin (*μg/g*)	TIMP-1 (*ng/g*)	Calprotectin (*μg/g*)	CEA (*μg/g*)
32.1	**<0.2**	**<0.2**	**0.9**	**24.4**	**26.3**
32.2	<0.2	<0.2	0.6	22.4	27.9
18.7	<0.2	<0.2	0.0	21.5	21.7
54.7	<0.2	<0.2	1.6	26.9	27.8
25.6	<0.2	<0.2	2.2	23.0	19.0
22.5	<0.2	<0.2	2.0	26.2	22.9
55.2	<0.2	<0.2	2.3	27.2	24.2
2153.4	**1.4**	**3.2**	**92.9**	**420.5**	**51.7**
393.8	0.2	0.2	22.2	179.2	51.7
2813.3	0.5	2.1	126.1	550.2	40.0
3746.9	3.6	5.8	107.6	542.5	49.7
2543.6	3.4	6.2	136.6	312.8	54.5
2314.7	3.9	7.0	123.3	677.7	56.8
2144.5	**4.9**	**6.2**	**93.5**	**350.3**	**42.9**

gen (CEA) (Roche Diagnostics GmbH, Mannheim, Germany). The first 3 assays were used as recommended by the manufacturer; however, an optimized stool extraction procedure was used. The extraction ratio (1:50) and the extract dilution factor (1:10) of these 3 assays was maintained. TIMP-1 and CEA were adapted to stool using 1:6 and 1:400 extract dilutions, respectively. An enzyme-linked immunosorbent assay was developed in-house for the detection of S100A12. Rabbits were immunized with recombinant full-length S100A12 expressed in *Escherichia coli*. The immunoglobulin G fraction was biotinylated or digoxigenylated to build a sandwich enzyme-linked immunosorbent assay using streptavidin-coated plates. Stool extracts were diluted 1:25 in sample dilution buffer and 50 μL of the diluted sample (or standard) was transferred to each well. Subsequently, 50 μL of antibody mix was added containing 0.5 μg/mL biotinylated polyclonal antibody "S100A12" and 0.5 μg/mL digoxigenylated polyclonal antibody "S100A12" in assay buffer. The plates were incubated for 60 minutes, washed 3 times with 350 μL of washing buffer, and 100 μL of 25 mU/mL digoxigenin-POD conjugate was added and incubated again for 60 minutes. Plates were washed 3 times with 350 μL of washing buffer, 100 μL ABTS solution was added, and then incubated for 60 minutes. The absorbance was measured at 405/620 nm. Recombinant full-length S100A12 was used for calibration.

Fecal Analysis

All stool samples were processed in a single laboratory with a modification of a recent procedure[10] using a freshly prepared extraction buffer (Tris 0.1 mol/L, pH 8.0, citric acid 0.1 mol/L, urea 1.0 mol/L, $CaCl_2$ 0.01 mol/L, bovine serum albumin 0.5%), adding a protease inhibitor cocktail (Mini Complete EDTA–free, Roche Diagnostics GmbH). The stool samples were thawed and 50 to 100 mg of each sample were transferred to a fecal sample preparation kit (Roche Diagnostics GmbH, Mannheim). A 50-fold excess of extraction buffer was added by weight, the samples were mixed vigorously for 30 minutes, transferred to a 10-mL tube, and centrifuged at 1200 × g for 10 minutes. Supernatants were filtered using a 5-μm cut-off filter (Ultrafree-CL, Millipore, Schwalbach, Germany), aliquoted, and stored at −70°C. The stool extracts were randomized and all biomarkers were measured in our laboratory with the exception of calprotectin, which was determined at an external site (Ärzte für Labormedizin Limbach und Kollegen, Heidelberg, Germany). Each marker was assessed independently from the same

Table 2. Marker Stability in Stool Extracts

Marker	Positive samples (*n*)	Mean recovery (±SD) after 1 day at RT (%)	Minimum recovery after 1 day at RT (%)	Mean recovery (±SD) after 3 days at RT (%)
S100A12	20	87 ± 15	52	73 ± 20
Hemoglobin	18	79 ± 23	45	59 ± 30
Hemoglobin-haptoglobin	15	78 ± 33	33	72 ± 30
TIMP-1	7	97 ± 32[a]	18[a]	100 ± 44[a]
Calprotectin	20	96 ± 10	74	94 ± 19
CEA	20	99 ± 6	95	100 ± 5

Table 3. Univariate Results

Marker	Median AUC[a] (p5–p95)	Sensitivity at 95% specificity, %		
		CRC collective I[b]	CRC collective II[c]	Advanced adenoma[c]
S100A12	0.95 (0.90–0.98)	82	84	13
Hemoglobin	0.91 (0.85–0.95)	82	87	20
Hemoglobin-haptoglobin	0.92 (0.88–0.97)	82	85	20
TIMP-1	0.92 (0.87–0.96)	73	72	10
Calprotectin	0.90 (0.84–0.95)	62	56	12
CEA	0.66 (0.57–0.73)	21	20	8

AUC, area under the curve of a receiver operating characteristics graph.
[a]One hundred–fold Monte Carlo cross-validation, median plus >0.05 and 0.95 quantile.
[b]One hundred–fold Monte Carlo cross-validation, median.
[c]Sensitivities were estimated in predictions (rule generated with CRC collective I vs controls).

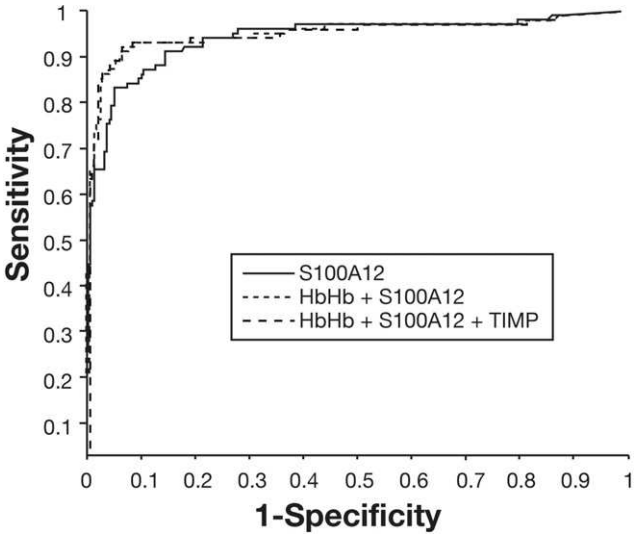

Figure 2. Receiver-operating characteristic curves of S100A12 and marker combinations in stool. The markers have been determined in CRC-collective I (101 patients) and controls (252 patients).

stool extract in duplicate. Laboratory personnel were unaware of clinical data. The maximum biomarker concentration of the respective sample pairs was used for further analysis.

Statistical Analysis

The diagnostic potential of the biomarkers was evaluated by receiver operator characteristics curves[11] and by determining the sensitivity at a preset specificity of 95% or 98%, respectively. Bayes Logistic Regression (BLR) was used as a mathematic model for marker combinations[12] as implemented in the Bayesian binary regression software. Results of the BLR were evaluated by 100 runs in a Monte-Carlo cross-validation design[13] applied on CRC collective I and controls. Within each run two thirds of all cases and controls, respectively, were selected randomly as a training set. BLR was applied to the training set to generate a diagnostic rule. A threshold on the estimated posterior case probabilities was determined on the controls of the training set to achieve an apparent specificity of 95% or 98% for the multivariate diagnostic rule. This rule then was applied to the remaining third of the data to estimate sensitivity and specificity at the given threshold. All multivariate results on the CRC collective I are therefore reported as median sensitivities from cross-validation. BLR then was applied to all samples in collective I to learn a final diagnostic rule and again its thresholds were determined. This rule with these thresholds then was applied to subgroups of the CRC collective I (UICC stages, left and right colon) and the apparent sensitivities are reported to see trends in stages and location of the cancer. Note that these

Table 4. Sensitivities of Marker Combinations

Collective	Patients	Sensitivity (%) at a specificity of 95%				
		Hemoglobin	Hemoglobin-haptoglobin	S100A12	S100A12 + hemoglobin-haptoglobin	S100A12 + hemoglobin-haptoglobin + TIMP-1
Median sensitivities from cross-validation						
CRC collective I	101[a]	82	82	82	88	88
Apparent sensitivities when applying the final multivariate rule to subcollectives						
UICC 0/I	23	74	78	57	78	78
UICC II	27	85	81	96	93	96
UICC III	12	92	83	92	92	83
UICC IV	23	83	83	87	96	96
Left-sided CRC	70	83	81	84	90	89
Right-sided CRC	31	81	81	81	84	87
CRC collective II	85	87	85	84	88	88
Advanced adenomas	113	20	20	13	22	20

estimates by construction can be overoptimistic, but the trends still give important information. The final rule also was used to predict test results for the CRC collective II and the adenomas. The apparent sensitivities in these groups were estimated without any overfit caused by learning.

Results

Marker Candidates

Six marker candidates were evaluated alone or in combination for the detection of CRC in stool samples: hemoglobin (iFOBT), hemoglobin-haptoglobin, calprotectin, CEA, TIMP-1, and S100A12.

Analyte Stability in Stool Extracts

The analyte stability was determined in stool extracts after storage at room temperature for 1 or 3 days, respectively. CEA (99% and 100%), calprotectin (96% and 94%), and TIMP-1 (97% and 100%) were very stable, followed by S100A12 (87% and 73%). Hemoglobin and hemoglobin-haptoglobin appeared to be less stable (Table 2). The interpatient variability of analyte recovery for hemoglobin and hemoglobin-haptoglobin was higher than for the other biomarkers.

Analyte Concentrations

All biomarkers were measured in the patient collectives described in Table 1. The median level of S100A12 was comparable in both CRC collectives (2153 vs 2145 ng/g), being much lower in the control group (32.1 ng/g), but without striking differences between the control subgroups (Table 1). The levels in the advanced adenoma collective were comparable with the concentrations found in the diverticulosis group (55.2 vs 54.7 ng/g), and only twice as high as in other controls. Similar results were found for TIMP-1 and calprotectin. With hemoglobin and hemoglobin-haptoglobin the median level in the control group was below the measuring range of the assays. In both assays the median was higher in CRC collective II than in collective I owing to the different inclusion criteria.

Univariate Analysis

The diagnostic performance for distinguishing CRC from controls was determined by receiver operator characteristic curve analysis. S100A12 showed the best discrimination followed by TIMP-1, hemoglobin-haptoglobin, hemoglobin, and calprotectin, whereas CEA did not reach a diagnostic relevant discriminatory power (Table 3). More important for a screening marker is the sensitivity at a high specificity level, which was arbitrarily set to 95%. S100A12, hemoglobin, and hemoglobin-haptoglobin all achieved a high sensitivity of 82% in CRC collective I (Table 3). The detection rate of advanced adenoma was low for all markers.

Multivariate Analysis

To test if marker combinations can improve the clinical performance we combined the markers using BLR. As shown in Figure 2, the area under the curve of S100A12 could be increased further by marker combinations (from 0.95 to 0.96). More important is the improvement of the sensitivity with the best single markers from 82% to 88% by combining S100A12 and hemoglobin-haptoglobin at 95% specificity (Table 4). A combination of more than 2 markers did not increase the sensitivity further. Although achieving a high sensitivity is of prime importance for CRC patients, a high specificity also is crucial for a screening marker to avoid distress by false-positive results. Hence, we determined the sensitivity of marker combinations at the even more restrictive specificity of 98%. The best sensitivity could be achieved using a combination of hemoglobin-haptoglobin, S100A12, and TIMP-1 (Table 4). Especially in early cancer stages UICC I and II, a strong increase was seen from 57% to 74% and from 74% to 93%, respectively, comparing univariate and multivariate analysis. Left-

Table 4. Continued

Sensitivity (%) at a specificity of 98%				
Hemoglobin	Hemoglobin-haptoglobin	S100A12	S100A12 + hemoglobin-haptoglobin	S100A12 + hemoglobin-haptoglobin + TIMP-1
70	73	67	79	82
52	57	30	52	74
70	74	81	81	93
75	75	75	83	83
78	78	70	87	91
73	71	67	76	86
65	74	61	77	84
79	78	66	82	86
12	12	4	9	12

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.