EVALUATION OF A 5-MARKER BLOOD TEST FOR COLORECTAL CANCER EARLY DETECTION IN A COLORECTAL CANCER SCREENING SETTING

Authors: Simone Werner1, Friedemann Krause2, Vinzent Rolny2, Matthias Strobl2, David Morgenstern3, Christian Datz4, Hongda Chen1, Hermann Brenner1,5

Affiliations: 1Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; 2Roche Diagnostics GmbH, Penzberg, Germany; 3Roche Diagnostic Operations, Inc, Indianapolis, IN, USA; 4Department of Internal Medicine, KH Oberndorf, Teaching Hospital of the Paracelsus Private Medical University Salzburg, Oberndorf, Austria; 5German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

Running title: A 5-Marker Blood Test for CRC Early Detection

Keywords: colorectal cancer, screening, multiple marker blood test, CEA, anti-p53

Financial support: This study was financed by Roche Diagnostics GmbH, Penzberg, Germany

Address for correspondence: Hermann Brenner, MD, MPH, Division of Clinical Epidemiology and Aging Research (C070), German Cancer Research Center, Im Neuenheimer Feld 581, D-69120 Heidelberg, Germany, Phone +49-6221-421300, fax +49-6221-421302, E-mail h.brenner@dkfz.de

Conflict of interest: This study was financed by Roche Diagnostics GmbH, Penzberg, Germany

Word count: abstract: 249, main-text: 4052 (without tables and figures)

Number of figures and tables: 3 figures, 3 tables, 2 supplementary tables
Statement of translational relevance

Blood tests for colorectal cancer (CRC) early detection would be a highly attractive alternative to endoscopical examinations and stool tests. In the last decade many candidate biomarkers were identified in studies that included symptomatic CRC cases recruited in clinics. However, the majority of these biomarkers failed in subsequent studies under screening conditions which stresses the importance of the study setting for biomarker discovery and validation. Here, we tested 1660 blood samples from participants of screening colonoscopy with a 5-marker blood test (carcinoembryonic antigen (CEA) + anti-p53 + osteopontin + seprase + ferritin). The diagnostic performance of the 5-marker test was comparable to guaiac-based fecal occult blood test but inferior to fecal immunological test. Of note, a combination of anti-p53 and CEA was sufficient to reach the same diagnostic performance under screening conditions as the whole 5-marker panel suggesting preference for these two markers for future multi-marker-panel development.
Abstract

Purpose: In initial studies that included colorectal cancer (CRC) patients undergoing diagnostic colonoscopy we had identified a serum marker combination able to detect CRC with similar diagnostic performance as fecal immunological test (FIT). In this study we aimed to validate the results in participants of a large CRC screening study conducted in the average-risk, asymptomatic screening population.

Experimental Design: We tested serum samples from 1200 controls, 420 advanced adenoma patients, 4 carcinoma in situ patients and 36 CRC patients with a 5-marker blood test (carcinoembryonic antigen (CEA)+anti-p53+osteopontin+seprase+ferritin). The diagnostic performance of individual markers and marker combinations was assessed and compared with stool test results.

Results: AUCs for the detection of CRC and advanced adenomas with the 5-marker blood test were 0.78 (95%CI: 0.68-0.87) and 0.56 (95%CI: 0.53-0.59), respectively, which now is comparable with guaiac-based fecal occult blood test (gFOBT) but inferior to FIT. With cutoffs yielding specificities of 80%, 90% and 95% the sensitivities for the detection of CRC were 64%, 50% and 42% and early-stage cancers were detected as well as late-stage cancers. For osteopontin, seprase and ferritin the diagnostic performance in the screening setting was reduced compared to previous studies in diagnostic settings while CEA and anti-p53 showed similar diagnostic performance in both settings.

Conclusions: Performance of the 5-marker blood test under screening conditions is inferior to FIT even though it is still comparable to the performance of gFOBT. CEA and anti-p53 could contribute to the development of a multiple marker blood-based test for early detection of CRC.
Introduction

With over 1.3 million new cancer cases worldwide and almost 700,000 deaths each year, colorectal cancer (CRC) is one of the most common cancers (1). Because of the slow progression from detectable and curable precancerous lesions to CRC and the strong dependence of prognosis on stage at diagnosis, early detection of CRC has great potential to reduce the burden of this disease (2-4). Colonoscopy is the gold standard for detecting CRC and its precursors, but its application as primary screening test is impaired by high costs, limited capacities and typically lower adherence (5, 6). Noninvasive stool tests are an attractive alternative for CRC screening due to their low cost and suitability for home use or in a primary care setting. However, the aversion to stool handling for guaiac-based fecal occult blood test (gFOBT) and fecal immunological tests (FITs) may limit the acceptance of stool tests (7). Blood-based tests may be more readily accepted by patients for many medical conditions and surveys conducted in individuals eligible for CRC screening showed a strong preference for collection of blood over feces (8). While the development of a simple and convenient blood test with diagnostic performance comparable to FIT could contribute to improvement in the acceptance of CRC screening, the diagnostic performance of currently existing blood tests is insufficient for their routine use (9).

The identification of a blood-based multi-marker panel for detection of CRC in a Marker Identification Program (MIP) study has been previously described (10). While most controls were recruited in a screening setting, CRC cases recruited from surgery centers were used to enrich the number of cancer cases due to the low prevalence of CRC in the screening population (10, 11). A multivariate analysis of 60 candidate biomarkers identified a 6-marker panel consisting of carcinoembryonic antigen (CEA), ferritin, seprase, osteopontin, anti-p53 autoantibody and CYFRA-21-1. With sensitivity and specificity of 69.6% and 95.0%,
respectively, for the detection of CRC, the diagnostic performance of the panel was comparable to the performance of FIT (10).

In order to validate this result in the target population of screening, we combined five of the six candidate biomarkers to a 5-marker blood test and retrospectively analyzed 1660 prospectively collected blood samples from all available CRC cases, most advanced adenoma cases and a selection of controls of the BliTz study collective. The BliTz study is an ideal setting for such a validation, as both cases and controls were recruited prior to diagnosis at screening colonoscopy (12-14). We aimed to determine the diagnostic performance of the 5-marker blood test for detection of CRC and advanced adenomas in the screening setting. Furthermore, we intended to compare the results with previous results from the diagnostic setting and with results of a commercial FIT and a commercial gFOBT.
Materials and Methods

Study design and sample selection

Samples were drawn from the BliTz study collective. BliTz (Begleitende Evaluierung innovativer Testverfahren zur Darmkrebs-Früherkennung) is an ongoing prospective screening study conducted in cooperation with more than 20 gastrointestinal practices in southwestern Germany. Detailed information on the Blitz study can be found elsewhere (12-14). Briefly, since the end of 2005 pre-colonoscopy stool- and blood samples from more than 7000 participants of screening colonoscopy were collected. After colonoscopy, basic demographic and clinical data were extracted from colonoscopy and histology reports in a standardized manner by trained research assistants who were blinded to blood and stool test results. Cancer stages were classified according to the UICC classification and advanced adenomas were defined as adenomas with at least one of the following features: 1cm in size, tubulovillous or villous components, high-grade dysplasia. Further patient data were collected with a short patient questionnaire. Personal and clinical data were stored separately in access databases for reasons of privacy protection. The study was approved by the Ethics Committee of the University of Heidelberg and informed consent was obtained from all participants. Auditing was conducted by the coordinating study center at the German Cancer Research Center and by the Coordination Centre for Clinical Trials (KKS) Heidelberg.

For the current analysis, participants with CRC (N=36), participants with carcinoma in situ (Cis, N=4) and participants with advanced adenoma (N=420) were compared with controls without colorectal neoplasms (N=1200). While the CRC and the Cis group were comprised of all available CRC/Cis patients from the BliTz study recruited until the 20th of February 2013, for the advanced adenoma and the control group representative samples were selected from the available BliTz participants. BliTz participants whose blood was taken after colonoscopy or for whom the date of blood withdrawal was unknown were excluded. Furthermore, study
participants with a history of CRC or inflammatory bowel disease, participants with a previous colonoscopy within the last 5 years and persons aged below 50 or over 79 were excluded, because these participants would typically not be considered to be the target population for CRC screening. In the control group also participants with insufficient bowel preparation before colonoscopy or incomplete colonoscopy were excluded to avoid false-negative colonoscopy results. Because this study was conducted in a true screening population in which CRC and adenoma patients are expected to be on average slightly older and to include a larger proportion of men, we did not match for these factors as this might lead to biased estimates of specificity in such a setting (15).

Handling of blood samples

After blood withdrawal in the gastrointestinal practices, serum samples were incubated at room temperature for 30-60 minutes to allow blood clotting and centrifuged at 2000-2500 g for ten minutes. Then they were transported to the DKFZ laboratory in a cold chain (medium transport time: 1 day), centrifuged again, aliquoted and stored at -80°C. For testing, the serum samples were randomized and shipped on dry ice to Roche Diagnostics GmbH, Penzberg, Germany. There was not more than one freeze-thaw cycle for each sample. The laboratory staff was blind to any information regarding the study population.

Immunoassays

The five biomarkers carcinoembryogenic antigen (CEA), ferritin, seprase, osteopontin and anti-p53 antibody were measured quantitatively on a cobas® e601 platform. The assays for each marker are designed as sandwich assays based on the Streptavidin-Biotin-technology. The capture antibodies are biotinylated and bind to streptavidin coated microparticles. The
secondary antibodies, covalently linked to Ruthenium complexes, are used for
electrochemoluminescent detection (16).

For CEA and ferritin the commercial tests Elecsys® CEA (Roche Diagnostics GmbH,
Catalogue Number: 11731629 322) and Elecsys® Ferritin (Roche Diagnostics GmbH,
Catalogue Number: 04491785 190) were used according to the instructions of the supplier.
Calibration was performed with the CEA CalSet (Roche Diagnostics GmbH, Catalogue
Number 11731645322) and the Ferritin CalSet (Roche Diagnostics GmbH, Catalogue
Number 03737586 190) in accordance with the package inserts.

Reagents for the quantitative analysis of seprase, osteopontin and anti-p53 antibody for the
cobas®e platform were available as prototypes at Roche Diagnostics. Performance
characteristics for prototype assays are expected to be similar to those seen for commercial
Elecsys® assays, i.e., repeatability CV <= 5-6%, intermediate precision/ total imprecision CV
<= 7-8%, dilution linearity within +/- 10%. For the prototype calibrators M-Cal-Seprase, M-
Cal-Osteopontin and M-Cal-anti p53 a full calibration was performed.

\textit{gFOBTs and FITs}

In the context of the BliTz study different stool tests for the early detection of CRC were
evaluated. For all except two participants included into this study FIT results were available
and for most study participants also gFOBT (HemOccult, Beckman Coulter GmbH, Krefeld,
Germany) results were available. While participants recruited before February 2009 were
tested by the quantitative FIT RIDASCREEN Hemoglobin (R-Biopharm AG, Bensheim,
Germany) as described elsewhere (17), for participants recruited since February 2009 the
quantitative FIT FOB Gold (Sentinel Diagnostics, Mailand, Italy) was used. Until January
2012 BliTz participants collected native stool samples that were immediately frozen and
thawed once before conducting the FIT according to the instructions of the manufacturers in a central laboratory. Since the end of January 2012, participants directly used the buffer-filled stool collection tubes from Sentinel for sample collection and mailed them to the central laboratory for analysis according to the instructions of the manufacturers. For all 3 FIT conditions (frozen stool + RIDASCREEN, frozen stool + FOB Gold, fresh stool in buffer-filled sentinel tubes + FOB Gold) we calculated cut-offs for test positivity based on all available BliTz controls with this FIT condition. At 96% specificity cut-offs were 9.6 µg hemoglobin/ g stool for the RIDASCREEN test, 42.5 µg hemoglobin/ g stool for the FOB Gold test with frozen stool samples and 15.3 µg hemoglobin/ g stool for the FOB Gold test with fresh stool samples collected in buffer-filled Sentinel tubes.

Data processing and Statistical analysis

For processing of the data obtained from the immunoassays the evaluation software OASE was used. For further data analyses statistical software (R version 3.1.0 (18), SAS version 9.2 (SAS Institute, Cary, NC, USA)) was used.

Basic demographic characteristics in the study population (sex, age, UICC stage) were summarized. The results of the five individual tests were combined into one single diagnostic result (the "score") at the biostatistics department of Roche Diagnostics using a defined algorithm with a pre-defined cutoff. This algorithm was selected by penalized LASSO regression on data of the MIP study (10) and re-optimized in a second large panel with screening controls and enriched cases (CT study). For the re-optimized algorithm, the MIP study marker CYFRA-21 was dispensable. That’s why this marker was not tested in the BliTz study anymore. An algorithm for the marker combination CEA + anti-p53 was derived from a
logistic regression model trained on data from the CT study. Socio-demographic
c characteristics of this study collective can be found in Supplementary Table 1.

Univariate marker results and results for marker combinations were compared between
participants with CRC and controls and between participants with advanced adenomas and
controls. Clinical performance (sensitivity and specificity with exact 95% confidence intervals
(CI)) and Receiver Operating Characteristic (ROC) curves for detection of CRC and advanced
adenoma were determined. In addition to analyses in the whole study population, we
performed stage-specific analyses. Areas under the ROC curves (AUCs) were compared by
the DeLong method with the R package “pROC” (19). For patients with gFOBT and FIT
results we also evaluated agreement between the 5-marker blood test and the stool tests.
Results

Characterization of the study population

For validation of the 5-marker blood test, 1660 participants (36 patients with invasive CRC, 4 Cis patients, 420 advanced adenoma patients and 1200 participants free of neoplasms) were selected from eligible participants of the CRC screening study BliTz as described in the Standards for the Reporting of Diagnostic accuracy studies (STARD) diagram (see Figure 1). Socio-demographic characteristics of all participants with valid measurement results (n=1656) are summarized in Table 1. As expected for a true colorectal cancer screening setting, the average age among CRC, Cis and advanced adenoma cases is slightly higher than among controls (mean ± standard deviation: 66.0 ± 6.2, 63.0 ± 5.3 and 63.6 ± 6.7 versus 62.0 ± 6.1 years). Also the proportion of men is higher among cases than among controls (CRC: 72.2% men, Cis: 75.0% men, advanced adenomas: 64.7% men, controls: 45.5% men). Among the group of CRC patients early-stage cancers (UICC stage I/II) were equally common as late-stage cancers (UICC stage III/IV).

Diagnostic performance of the original marker panel

For blood samples from all except four subjects (one advanced adenoma patient and three controls) valid measurements for all five Elecsys® Assays (CEA, ferritin, seprase, osteopontin and anti-p53) could be obtained. We used a predefined algorithm obtained in the MIP study and validated in the CT study to combine the results from the five Elecsys® Assays into a single prediction score. ROC curve analysis revealed an AUC of 0.78 (95% CI: 0.68-0.87) for the discrimination of CRC patients (Stage I-IV) and controls and an AUC of 0.56 (95% CI: 0.53-0.59) for the discrimination of advanced adenoma patients and controls (see Figure 2). In a sensitivity analysis including the four Cis patients the diagnostic
performance for the detection of CRC (stage 0-IV) was slightly worse with an AUC of 0.76 (95% CI: 0.67-0.85).

When using a cutoff at 90% specificity in the CT study, the 5-marker combination yielded sensitivities of 44 (95% CI: 28-62)% and 12 (95% CI: 9-15)% for CRC and advanced adenomas at a specificity of 94 (95% CI: 92-95)%. When adjusting the cutoffs to yield specificities of 80%, 90% and 95% in the BliTz collective, sensitivities for the detection of CRC were 64%, 50% and 42%. These values are lower than the sensitivities observed in the MIP study (10) and the CT study (see Supplementary Table 2). Sensitivities for advanced adenomas were below 30% even at cutoffs yielding 80% specificity (see Table 2).

With AUCs of 0.80 (95% CI: 0.67-0.93) and 0.75 (95% CI: 0.62-0.88) the ability of the 5-marker blood test to detect early and late-stage cancer was similar (p-value: 0.57) (see Table 2).

Diagnostic performance of single markers

To further evaluate the loss of performance of the 5-marker panel in the BliTz study collective, we compared the univariate results for CEA, ferritin, seprase, osteopontin and anti-p53 in the BliTz and in the CT study. Interestingly, two of the five markers (CEA and anti-p53) showed very similar diagnostic performance in both studies (AUC in BliTz: 0.84 (95% CI: 0.78-0.90) and 0.57 (95% CI: 0.51-0.63); AUC in CT study: 0.77 (95% CI: 0.73-0.81) and 0.59 (95% CI: 0.56-0.62)) while the other markers didn’t (see Figure 3a-e). The largest decrease of diagnostic performance was seen for seprase with an AUC of 0.78 (95% CI: 0.74-0.82) in the CT study and an AUC of 0.60 (95% CI: 0.49-0.70) in the BliTz study.
Diagnostic performance of a 2-marker combination

Because of the poor univariate results for seprase, osteopontin and ferritin in the BliTz study collective, we decided to perform an exploratory evaluation of the marker combination CEA + anti-p53. For training of the algorithm, data from the CT study were used. In the BliTz study the 2-marker combination reached an AUC of 0.85 (95% CI: 0.78-0.91) for the detection of CRC (see Figure 3f) and an AUC of 0.56 (95% CI: 0.53-0.59) for the detection of advanced adenomas. When adjusting the specificity to 80%, 90% and 95%, the sensitivities for CRC detection were 67%, 58%, and 47% for the 2-marker combination. For the same specificities, the sensitivities of CEA alone were 62%, 50% and 40%.

Comparison of the 5-marker blood test with stool tests

For 23 CRC cases, 301 advanced adenoma cases and 899 controls, results of the 5-marker blood test and both stool tests (gFOBT and FIT) were available. Table 3 shows the diagnostic performance of all tests in this subpopulation. For reasons of comparison, the specificities of the FIT and the 5-marker blood test were adjusted to the specificity of the gFOBT (96%). At this specificity, both the 5-marker blood test and the gFOBT could identify 39% of all CRC cases. Of note, the gFOBT and the 5-marker blood test primarily detected different CRC patients which led to an increased sensitivity of 65% (at minimal loss of specificity) when combining gFOBT and blood test results. For advanced adenomas the sensitivities of the 5-marker blood test, the gFOBT and the combination of both tests were 8%, 4% and 12%. With sensitivities of 78% for CRC and 28% for advanced adenomas the FIT was superior to the other tests and its combination with the 5-marker panel could only slightly further increase sensitivities (83% for CRC, 35% for advanced adenomas) at the cost of a lower specificity (92%).
Discussion

In this study we evaluated a 5-marker blood test for colorectal cancer early detection in a large screening study. The AUCs for the detection of CRC and advanced adenomas were 0.78 (95% CI: 0.68-0.87) and 0.56 (95% CI: 0.53-0.59), respectively. At specificities of 80%, 90% and 95% the sensitivities for the detection of CRC yielded 64%, 50% and 42% and early-stage cancers were detected at least as well as late-stage cancers. Of note, two of the five markers (CEA and anti-p53) were sufficient to achieve a similar or even better diagnostic performance for CRC (AUC of 0.85 (95% CI: 0.78-0.91). In a subsample of participants with gFOBT, FIT and blood test results we directly compared the performance of these tests. With a sensitivity of 39% at a specificity of 96% gFOBT and the 5-marker blood test performed equally well, and sensitivity increased to 65% when both tests were combined. Nevertheless, both tests and their combination were outperformed by FIT.

In our previous MIP study a six marker blood test (5-marker panel with the addition of CYFRA21-1) was able to detect CRC patients with a sensitivity of 69.6% at a specificity of 95%, which was similar to the diagnostic performance of FIT in the MIP study (10). In the CT study, that was used to refine the algorithm for the 5-marker blood test, the observed sensitivity of the 5-marker blood test was 68.1% at 95% specificity. The obvious drop of diagnostic performance of the 5-marker blood test in participants selected from the BliTz study collective may be explained by differences in the study populations. While BliTz is a real screening study in which all cases were recruited before diagnosis at screening colonoscopy, CRC cases in the MIP and CT study had to be enriched with patients recruited at surgery units. For clinically recruited CRC cases there is the possibility that blood marker levels are altered by diagnostic or therapeutic interventions or lifestyle changes in response to the CRC diagnosis. Furthermore, in contrast to screening settings, in clinical settings it can’t be guaranteed that blood withdrawal and blood storage conditions at recruitment site are
exactly the same for cases and controls. Last but not least, CRC patients recruited in clinical
settings are often in a more advanced stage and more often present symptoms than CRC
patients diagnosed at screening colonoscopy (20).

Our univariate analyses suggest that not all markers of the 5-marker blood test are as prone to
study setting issues as others. With an AUC of 0.84 the diagnostic performance of CEA was
even better than in the CT study (AUC: 0.77) and for anti-p53 the AUCs in both studies were
quite similar (AUC in BliTz: 0.57, AUC in CT study: 0.59). These results suggest that the
observed diagnostic performance for CEA and anti-p53 indeed represents true cancer-specific
differences between cases and controls. For ferritin, osteopontin and seprase other factors
might have contributed to the previously observed good diagnostic performance for CRC
detection. Serum ferritin levels, for instance, can be influenced by age, sex, body mass index,
acute or chronic inflammation and aspirin use in addition to cancer (21). For CRC the
situation is especially complex, because possible positive correlations between cancer-specific
processes and ferritin, as found for other cancers (21) might be antagonized by iron-
deficiency anemia caused by chronic gastrointestinal bleedings.

The sensitivity of anti-p53 is limited due to the fact that not all CRC patients have p53-
mutations and not all patients with p53 mutations produce antibodies against this tumor
suppressor protein (22). Nevertheless, the remarkably high specificity of anti-p53 for cancer,
which also can be seen in the very steep slope of the left part of its ROC curve (see Figure
3b), makes it possible to increase the sensitivity of conventional tumor markers without
reducing specificity to a relevant extent when anti-p53 is added in a marker combination (23).
The combined use of CEA and anti-p53 for CRC detection has been evaluated in earlier
studies and sensitivities between 33% and 73% have been reported (22-26). However, none of
these studies was performed in a screening setting and only Kojima at al. reported specificities
for this marker combination. In their newer and larger study from 2011 a sensitivity of 48% at
93% specificity for the detection of CRC was reached (25), which is very similar to our findings in participants of screening colonoscopy. It should be stated that both, CEA and anti-p53, are not cancer-type specific and have been found in patients with other cancers like lung cancer (27, 28). Thus there is the possibility that some of the controls with false positive test results actually are persons with an undiscovered other malignancy.

With 47% and 6% sensitivity at 95% specificity, the capability of the marker combination CEA + anti-p53 to detect CRC and advanced adenomas was comparable or even superior to Epi proColon®, to our knowledge the only prospectively evaluated blood test for early detection of CRC so far (29). Although the 2-marker combination CEA + anti-p53 alone cannot compete with the fecal immunological test, it appears plausible that a combination with further blood biomarkers, such as DNA methylation markers (30), microRNA markers (31), autoantibody markers (24) or protein markers (32, 33) might increase the diagnostic performance sufficiently for an application in mass screening.

Blood tests seem to be better accepted in public than stool tests. For instance in a study by Adler et al. over 100 persons that refused to participate in screening colonoscopy were offered a choice of a blood-based or a stool-based CRC early detection test and while 83% of the participants picked the blood test, only 15% picked the stool test (34). One major advantage of FIT over current blood tests, in addition to the better diagnostic performance for detection of CRC, is its ability to detect a relevant proportion of advanced adenomas. Stool tests might have a larger potential to capture localized tumor effects in general such as excretion of blood or components of tumor cells with stool which would be hard if not impossible to detect by blood tests, in particular at early tumor stage and for precursors of CRC. In our analyses the diagnostic performance for the detection of advanced adenomas was poor and the diagnostic performance for the detection of non-advanced adenomas is expected to be even worse. Non-advanced adenomas were deliberately not included in our analyses as their transition rates to
more advanced neoplasms are low (35) and there is an ongoing debate if they should be considered as one of the target lesions for CRC screening or not (36). For the development of future blood tests for early detection of CRC it would be beneficial to identify markers that also detect advanced adenomas. In the meantime efforts to increase public’s adherence for stool test should be enhanced.

To our knowledge, our study is the first to test a 5-marker blood test, including CEA and anti-p53, in subjects from a true screening setting. There are specific strengths and limitations that have to be considered. A strength is that cases and controls were selected from participants of screening colonoscopy which represent the target population for CRC screening. With over 1600 study participants, including 1200 controls and 400 advanced adenoma patients our sample size was very large. So estimates for specificity and sensitivity for the detection of advanced adenomas could be determined very precisely and 95% confidence intervals are small. Furthermore, we used a predefined algorithm trained on data of our previous studies, to avoid overfitting, a serious problem seen in many multi-marker studies (37). A limitation is the relatively small number of CRC cases included in this study which is due to the low prevalence of CRC in participants of screening colonoscopy. This limited our options to perform subgroup specific analyses. In addition, we did not evaluate the value of the 5-marker blood test or individual markers for prognosis or monitoring CRC patients.

In conclusion, the validation of a 5-marker blood test for CRC early detection in participants selected from a screening study collective uncovered decreased diagnostic performance for the markers ferritin, osteopontin and seprase, compared to previous evaluations in studies conducted among cases recruited in clinical settings. Thus the overall diagnostic performance estimates for the 5-marker blood test dropped from values comparable with FIT in the clinical setting to values comparable with gFOBT in our study. Our results furthermore underline the potential of CEA and anti-p53 to discriminate cancer patients and controls under screening.
conditions suggesting their potential to contribute to the development of a multiple marker blood-based test for early detection of CRC.

Acknowledgement

We gratefully acknowledge the excellent cooperation of gastroenterology practices in patient recruitment and of Labor Limbach in sample collection. We thank Dr. Katja Butterbach and Ulrike Schlesselmann for their excellent work in laboratory preparation of blood samples. We also thank Isabel Lerch, Susanne Köhler, Utz Benscheid, Jason Hochhaus and Maria Kuschel for their contribution in data collection, monitoring and documentation.
References


**Table 1.** Study population characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control</th>
<th>Advanced Adenoma</th>
<th>Carcinoma in situ</th>
<th>CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>1200</td>
<td>420</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Measurements available</td>
<td>1197</td>
<td>419</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Age [mean ± SD, years]</td>
<td>62.0 ± 6.1</td>
<td>63.6 ± 6.7</td>
<td>63.0 ± 5.3</td>
<td>66.0 ± 6.2</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male [n, %]</td>
<td>545 (45.5%)</td>
<td>271 (64.7%)</td>
<td>3 (75.0%)</td>
<td>26 (72.2%)</td>
</tr>
<tr>
<td>Female [n, %]</td>
<td>652 (54.5%)</td>
<td>148 (35.3%)</td>
<td>1 (25.0%)</td>
<td>10 (27.8%)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UICC 0</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>UICC I</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td>UICC II</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>UICC III</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>UICC IV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviations: SD = standard deviation, n = number.
Table 2. Sensitivities and specificities of the 5-marker blood test at different cutoffs.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Cutoff obtained in CT study</th>
<th>Cutoff adjusted at 95% spec</th>
<th>Cutoff adjusted at 90% spec</th>
<th>Cutoff adjusted at 80% spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal Cancer</td>
<td>36</td>
<td>44 [28-62]%</td>
<td>42 [26-59]%</td>
<td>50 [33-67]%</td>
<td>64 [46-79]%</td>
</tr>
<tr>
<td>UICC I-II</td>
<td>18</td>
<td>50 [26-74]%</td>
<td>44 [22-69]%</td>
<td>50 [26-74]%</td>
<td>72 [47-90]%</td>
</tr>
<tr>
<td>UICC III-IV</td>
<td>18</td>
<td>39 [17-64]%</td>
<td>39 [17-64]%</td>
<td>50 [26-74]%</td>
<td>56 [31-78]%</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td>4</td>
<td>0 [0-60]%</td>
<td>0 [0-60]%</td>
<td>25 [1-81]%</td>
<td>25 [1-81]%</td>
</tr>
<tr>
<td>Advanced Adenoma</td>
<td>419</td>
<td>12 [9-15]%</td>
<td>9 [6-12]%</td>
<td>16 [12-19]%</td>
<td>25 [21-30]%</td>
</tr>
</tbody>
</table>

Sensitivity of the 5-marker blood test [95% CI]

Specificity of the 5-marker blood test [95% CI]

Abbreviations: CI = confidence interval, spec = specificity
**Table 3.** Comparison of 5-marker blood test, gFOBT and FIT in participants with stool test results.

<table>
<thead>
<tr>
<th></th>
<th>5-marker blood test</th>
<th>gFOBT</th>
<th>5-marker blood test + gFOBT*</th>
<th>FIT</th>
<th>5-marker blood test + FIT*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensitivity CRC</strong> [95% CI]</td>
<td>39 [20 - 61]%</td>
<td>39 [20 - 61]%</td>
<td>65 [43 - 84]%</td>
<td>78 [56 - 93]%</td>
<td>83 [61-95]%</td>
</tr>
<tr>
<td><strong>Sensitivity advanced adenoma</strong> [95% CI]</td>
<td>8 [5 - 12]%</td>
<td>4 [2 - 7]%</td>
<td>12 [9 - 16]%</td>
<td>28 [23 - 34]%</td>
<td>35 [30-41]%</td>
</tr>
<tr>
<td><strong>Specificity</strong> [95% CI]</td>
<td>96 [94 - 97]%</td>
<td>96 [94 - 97]%</td>
<td>92 [90 - 93]%</td>
<td>96 [94 - 97]%</td>
<td>92 [90-94]%</td>
</tr>
</tbody>
</table>

* The combination 5-marker blood test + gFOBT/ FIT was considered positive if either the 5-marker blood test, the gFOBT/ FIT or both tests were positive. ** n = 23 *** n = 301 **** n = 899.

Abbreviations: CI = confidence interval.
Figure legends

**Figure 1.** STAndards for the Reporting of Diagnostic accuracy studies (STARD) diagram of participants of the BliTz study (December 2005 - February 2013). 1420 is the number of available advanced adenoma patients in BliTz that was anticipated at the time of study design. Preferentially, participants were selected from the subgroup with valid FIT results (FIT result available and stool sampling before colonoscopy) which comprises > 90% of all BliTz participants.

**Figure 2.** Diagnostic performance of the 5-marker blood test for the detection of CRC (a) and advanced adenomas (b).

**Figure 3.** (a)-(e) Univariate analysis: Performance of CEA, anti-p53, osteopontin, ferritin and seprase in the BliTz and in the CT study. (f) Diagnostic performance of the 2-marker combination anti-p53 + CEA in the BliTz study.
Participants with signed consent, colonoscopy results and questionnaires participating until February 20th, 2013 (N = 5781)

Exclusion of participants without adequate blood samples:
- No serum sample available (N = 332)
- Blood withdrawal after colonoscopy (N = 38) or unknown time of blood withdrawal (N = 158)

Exclusion of participants who do not represent the target population of screening:
- History of CRC or inflammatory bowel disease (N = 41)
- Colonoscopy history in the last 5 years or unknown colonoscopy history (N = 282 + 70)
- Age <50 years, age ≥80 years or age unknown (N = 89 + 58 + 1)

Exclusion of participants with potentially false negative results (only in participants free of neoplasms):
- Inadequate bowel preparation before colonoscopy (N = 307)
- Incomplete colonoscopy (N = 62)

Participants eligible for sample selection (N = 4567)

CRC and carcinoma in situ (N = 42)
- Not further defined polyp (N = 36)
- Advanced adenoma (N = 483)
- Nonadvanced adenoma (N = 916)
- Free of colorectal neoplasms (N = 3090)

Exclusion of participants with critical information missing at time of sample selection (N = 2)

Available CRC and carcinoma in situ patients (N = 40)

Representative sample of adenoma patients (N = 420)

Representative sample of controls free of neoplasms (N = 1200)

Sample set for measurements with the 5-marker blood test (N = 1660)

No measurement results (N = 4)

Samples with valid measurement results (N = 1656)

Figure 1. STAndards for the Reporting of Diagnostic accuracy studies (STARD) diagram of participants of the BliTz study (December 2005 - February 2013). 420 is the number of available advanced adenoma patients in BliTz that was anticipated at the time of study design. Preferentially, participants were selected from the subgroup with valid FIT results (FIT result available and stool sampling before colonoscopy) which comprises >90% of all BliTz participants.
Evaluation of a 5-Marker Blood Test for Colorectal Cancer Early Detection in a Colorectal Cancer Screening Setting

Simone Werner, Friedemann Krause, Vinzent Rolny, et al.

*Clin Cancer Res* Published OnlineFirst November 11, 2015.

**Updated version**
Access the most recent version of this article at:

**Supplementary Material**
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2015/11/11/1078-0432.CCR-15-1268.DC1

**Author Manuscript**
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

---

**E-mail alerts**
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.