

Development and validation of a panel of five proteins as blood biomarkers for early detection of colorectal cancer

Hongda Chen^{1,2}
Jing Qian¹
Simone Werner¹
Katarina Cuk¹
Phillip Knebel³
Hermann Brenner^{1,4,5}

¹Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; ²Program Office for Cancer Screening in Urban China, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People's Republic of China; ³Department of General, Visceral and Transplantation Surgery, University of Heidelberg; ⁴Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases; ⁵German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

Correspondence: Hermann Brenner
Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany
Tel +49 6221 42 1301
Fax +49 6221 42 1302
Email h.brenner@dkfz-heidelberg.de

Objective: Reliable noninvasive biomarkers for early detection of colorectal cancer (CRC) are highly desirable for efficient population-based screening with high adherence rates. We aimed to discover and validate blood-based protein markers for the early detection of CRC.

Patients and methods: A two-stage design with a discovery and a validation set was used. In the discovery phase, plasma levels of 92 protein markers and serum levels of TP53 autoantibody were measured in 226 clinically recruited CRC patients and 118 controls who were free of colorectal neoplasms at screening colonoscopy. An algorithm predicting the presence of CRC was derived by Lasso regression and validated in a validation set consisting of all available 41 patients with CRC and a representative sample of 106 participants with advanced adenomas and 107 controls free of neoplasm from a large screening colonoscopy cohort (N=6018). Receiver operating characteristic (ROC) analyses were conducted to evaluate the diagnostic performance of individual biomarkers and biomarker combinations.

Results: An algorithm based on growth differentiation factor 15 (GDF-15), amphiregulin (AREG), Fas antigen ligand (FasL), Fms-related tyrosine kinase 3 ligand (Flt3L) and TP53 autoantibody was constructed. In the validation set, the areas under the curves of this five-marker algorithm were 0.82 (95% CI, 0.74–0.90) for detecting CRC and 0.60 (95% CI, 0.52–0.69) for detecting advanced adenomas. At cutoffs yielding 90% specificity, the sensitivities (95% CI) for detecting CRC and advanced adenomas were 56.4% (38.4%–71.8%) and 22.0% (13.4%–35.4%), respectively. The five-marker panel showed similar diagnostic efficacy for the detection of early- and late-stage CRC.

Conclusion: The identified most promising biomarkers could contribute to the development of powerful blood-based tests for CRC screening in the future.

Keywords: amphiregulin, growth differentiation factor 15, adenoma, colorectal cancer, screening

Plain language summary

What is the current knowledge?

- Early detection of colorectal cancer (CRC) can strongly reduce the burden of this disease.
- Screening by fecal occult blood tests (FOBTs) has been shown to reduce CRC mortality in randomized controlled trials.
- Blood-based tests could be an important alternative or supplement to fecal tests for noninvasive screening with potentially better adherence.
- It is essential to identify biomarkers and evaluate their diagnostic performance in true screening settings.

What is new here?

- A five-marker algorithm was constructed and independently validated in prospective samples from a cohort of screening colonoscopy participants during 2005–2014 (N=7197).

- The five-marker algorithm showed comparable or even better diagnostic performance in detecting CRC and its precursors than plasma methylated Septin 9 and FOBT in external validations.

Introduction

With over 1.4 million new cancer cases and 693,900 deaths estimated to have occurred in 2012, colorectal cancer (CRC) is the third most commonly diagnosed cancer in males and the second in females.¹ Randomized controlled trials and observational studies have shown that screening by endoscopic examinations or stool-based tests yield a reduction in CRC incidence and mortality.²⁻⁵ However, the participation rates in endoscopy and stool-based screening offers are often relatively low. For example, in an ongoing randomized clinical trial (NordICC study) conducted in Poland, Norway, the Netherlands and Sweden, the participation rates of screening colonoscopy varied from 22.9% to 60.7%.⁶ The participation rates of fecal occult blood tests (FOBTs) varied from 7.0% to 67.7% across CRC screening programs worldwide.⁷ The relatively low compliance of endoscopic examinations and stool tests might limit the screening efficacy on a population level.

Blood tests might potentially improve adherence in population-based screening programs, given their minimally invasive nature and straightforward implementation in routine medical examinations.⁸ To date, plasma methylated Septin 9 is the first and only blood-based test approved by the US Food and Drug Administration (FDA) for CRC screening.⁹ However, the diagnostic performance of this test is not optimal, with sensitivity for detecting CRC and advanced adenomas of 48.2% and 11.2%, respectively, at 91.5% specificity.¹⁰ Other alleged promising blood biomarkers, such as autoantibodies¹¹ and microRNAs,¹² were rarely validated in the targeted screening populations.

In a previous study, we prospectively evaluated 92 plasma proteins as potential biomarkers in the early detection of CRC.¹³ Although promising results were reported, this study was limited by relatively small sample size (N=92) and the diagnostic potential of biomarkers for detecting CRC-related precursors was not evaluated.¹³ In the current study, we tested an updated panel of 92 tumor-associated proteins and TP53 autoantibody in a much larger sample. We aimed to discover individual biomarkers and define optimal multi-marker combinations for the early detection of CRC in a training set. Estimates of diagnostic performance of identified biomarkers and multi-marker panels for detecting CRC and its precursors were independently validated in prospectively selected samples exclusively recruited in a true screening setting.

Patients and methods

Study design and study population

In our analyses, a two-step approach with selection of biomarkers and construction of multi-marker algorithms in a discovery set and validation of the findings in an independent validation set was adopted. Samples from two study populations were used, including clinically detected CRC cases recruited at hospitals (used for marker discovery only) and participants with CRC or advanced adenomas, as well as control participants without colorectal neoplasms recruited in a true screening setting (BLITZ study).

The prediagnostic samples from the BLITZ study

The BLITZ study (Begleitende Evaluierung innovativer Testverfahren zur Darmkrebsfrüherkennung [Evaluation of innovative test method for early detection of colorectal cancer]) is an ongoing cohort of participants attending screening colonoscopy in Germany. Detailed information on the BLITZ study has been reported elsewhere.^{14,15} Briefly, this study is conducted in collaboration with 20 gastroenterology practices in southern Germany since November 2005. Participants are recruited at a preparatory visit at practices typically 1 week before screening colonoscopy and are invited to donate prediagnostic blood and stool samples. Self-administrated questionnaires regarding potential risk factors known or suspected to be related to CRC, such as family history, smoking, diet and physical activity information, are collected from all participants. The German screening colonoscopy program, introduced in October 2002, offers up to two screening colonoscopies at least 10 years apart to men and women aged 55 years or older.¹⁶ The high quality of screening colonoscopy in Germany is reflected in high adenoma detection rates which have steadily increased since the introduction of the screening program.¹⁷

By September 2014, a total number of 7,197 participants have been recruited (Figure 1). For this analysis, the following exclusion criteria were applied: 1) missing plasma samples; 2) blood taken after screening colonoscopy; 3) inflammatory bowel disease or previous CRC; 4) insufficient bowel preparation (only for individuals with no significant findings at screening colonoscopy); and 5) incomplete colonoscopy (only for individuals with no significant findings at screening colonoscopy). From the remaining participants of the BLITZ study recruited during 2005–2015 (N=6018), all 43 screening detected CRC cases, as well as random samples of individuals with advanced colorectal adenomas (N=113) or with no colorectal neoplasms (N=233), were included in this analysis. Because this study was conducted in a true screening

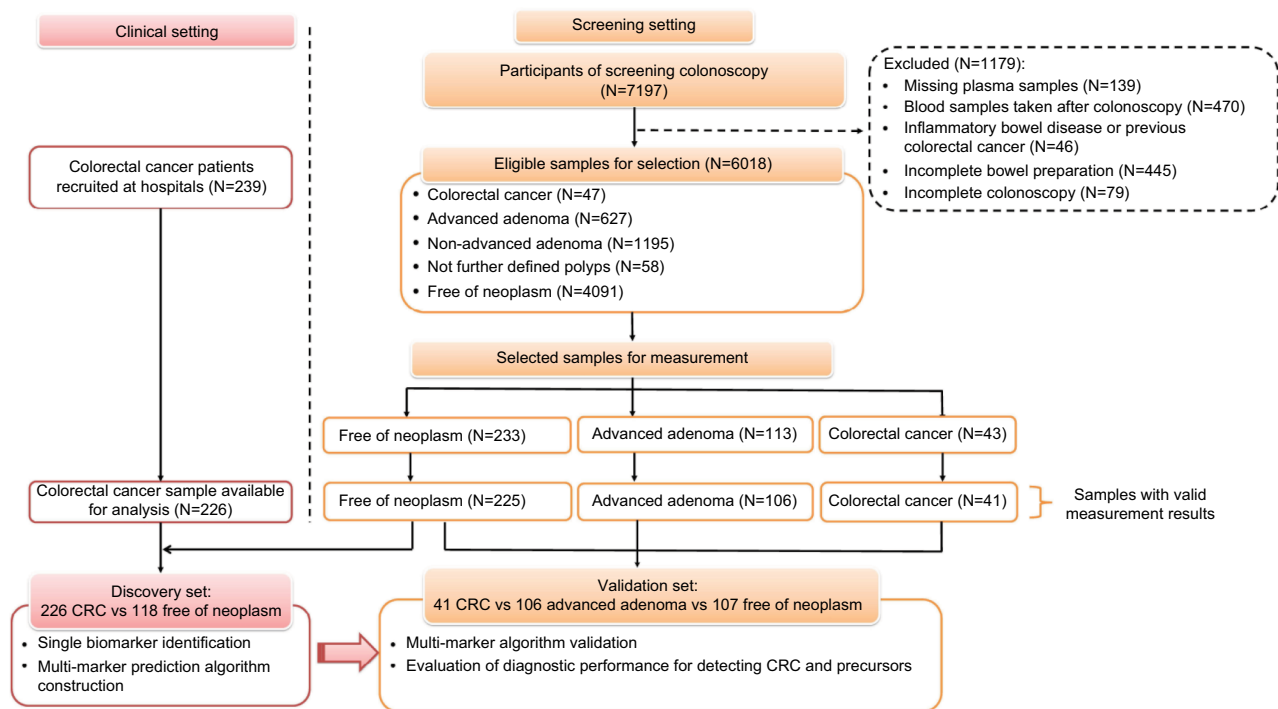


Figure 1 STARD diagram showing the selection of study participants enrolled in the BLITZ study during 2005–2015 and the analysis scheme.

Abbreviations: CRC, colorectal cancer; STARD, STAndards for the Reporting of Diagnostic accuracy studies.

population in which CRC patients are expected to be on average slightly older and to include a somewhat large proportion of men, we did not match for these factors as this might lead to biased estimates of specificity in such a setting.¹⁸

Clinically selected CRC samples

Given the limited number of CRC cases identified in the screening setting even in a study as large as BLITZ, we additionally included 239 CRC patients recruited after diagnosis but before treatment at four hospitals in southern Germany for the discovery phase of this study. These patients were different from the much smaller number of such patients ($n=54$) included in our previous study.¹³ The same questionnaire data and medical records were collected, and blood samples were obtained according to identical standard operating procedures (SOPs).

All studies were approved by the ethics committee of the Heidelberg University Medical School and the ethics committees of the state physicians' boards of Baden-Württemberg and Rhineland Palatinate. Written informed consent was obtained from each participant.

Classification of CRC and advanced adenomas

CRC stage was classified according to the Union for International Cancer Control (UICC) tumor-node-metastasis (TNM) stage classification (seventh version). Participants of screening

colonoscopy were classified according to the most advanced finding reported in the colonoscopy and/or histology report. Advanced adenomas were defined as adenomas with at least one of the following features: 1) high-grade dysplasia (HGD); 2) villous or tubular–villous architecture; and 3) size ≥ 10 mm. Relevant information was extracted from colonoscopy and hospital records by two research assistants independently who were blinded with respect to the blood test results.

Laboratory procedures

Sample preparation

Blood samples were collected in Sarstedt S-Monovette K3 EDTA or BD Vacutainer K3 EDTA tube prior to bowel preparation for bowel colonoscopy (BLITZ study), or prior to large bowel surgery or neoadjuvant therapy (239 CRC cases from the clinical setting). The samples were transported to the laboratory while preserving a cold chain and were centrifuged at $2000\text{--}2500\times g$ for 10 min, aliquoted and stored at -80°C until further use. Details on the SOPs have also been described previously.¹⁹

Protein profiling

A total of 92 predefined human tumor-associated protein biomarkers were measured in 628 samples using Proseek Multiplex Oncology I v2^{96x96} (Olink Bioscience, Uppsala, Sweden; the full protein marker list is provided in [Table S1](#)). The panel of 92 protein biomarkers reflects various biologic

mechanisms involved in carcinogenesis, such as angiogenesis, cell–cell signaling, growth control and inflammation. All laboratory operations were conducted according to the Proseek Multiplex Oncology I^{96×96} User Manual at Olink Bioscience. In short, the Proseek reagents are based on the proximity extension assay (PEA) technology, where 92 oligonucleotide-labeled antibody probe pairs are allowed to bind to their respective target present in the sample. A polymerase chain reaction (PCR) reporter sequence is formed by a proximity-dependent DNA polymerization event and is subsequently detected and quantified using real-time PCR. The laboratory operators were blinded with respect to any information regarding the study participants.

Antibodies against TP53

For the majority of study participants, including 239 cases with clinically identified CRC, 39 cases with screening-detected CRC, 82 participants with advanced adenomas and 208 controls free of neoplasm, measurements of antibodies against TP53 were available from a previous study, which has been reported elsewhere.²⁰ Briefly, antibodies against TP53 were measured by multiplex serology, a fluorescent bead-based glutathione S-transferase (GST) capture immunosorbent assay, as described previously.^{21,22}

Data normalization and statistical analyses

Data normalization of protein profiling

The raw data of the protein profiling were first normalized following the standard protocol from the manufacturer and using the Olink Wizard of GenEx software (MultiD, Göteborg, Sweden). For each data point, the raw quantification cycle value (Cq-value, in log₂ scale) was exported from the Fluidigm Real-Time PCR Analysis Software. The Cq-value is defined as the calculated cycle number at which the PCR product crosses a threshold of detection and is used to represent the expression levels of respective proteins in the current study. The first step of normalization was to subtract the raw Cq-value for the extension control for the corresponding sample to correct for technical variation. The calculated Cq-values (dCq-value) were further normalized against the negative control determined in the measurement, which yielded ddCq-values (hereafter, Cq-value, in log₂ scale) and could be used for further analyses. Limit of detection (LOD) was defined as the mean value of the three negative controls plus three calculated standard deviations. A total of 30 samples with invalid test results were excluded from this analysis. Missing data and data with a value lower than LOD were replaced with LOD in the following statistical analyses.

Statistical analyses

In our analysis, the discovery set samples included CRC cases recruited in the clinical setting and 118 randomly selected controls free of neoplasm from the BLITZ study (Figure 1). The validation set was defined in such a way that it represents a true screening setting, i.e., only participants from the BLITZ study were included.

The plasma protein levels (Cq-value) were first compared between CRC cases and neoplasm-free controls in the discovery set samples and validation set samples using Wilcoxon rank-sum test (hereafter, Wilcoxon test). The Benjamini–Hochberg method was additionally employed to correct for multiple testing.

A multi-marker algorithm was derived by applying the Lasso logistic regression model based on significant biomarkers identified in the discovery set samples. A second prediction algorithm was built by combining the measurements of the selected protein biomarkers from the Lasso logistic regression model with TP53 autoantibody measurements using logistic regression. Both prediction algorithms were further validated using receiver operating characteristic (ROC) curves in the validation set. Areas under the curve (AUCs) and sensitivities at cutoffs yielding 80% and 90% specificity, respectively, and their 95% CIs of the multi-marker algorithms were calculated and reported. In addition, we conducted subgroup analyses on the diagnostic performance of the multi-marker algorithms according to sex, age (<65 vs ≥65 years) and cancer stage in the validation set.

Statistical analyses were performed with the statistical software R version 3.0.3.²³ All tests were two-sided, and *p*-values of 0.05 or less were considered to be statistically significant.

Results

Figure 1 shows the STAndards for the Reporting of Diagnostic accuracy studies (STARD) diagram, showing the selection of study participants enrolled in the BLITZ study during 2005–2015 and also the scheme of analysis. The discovery set included 226 clinically recruited CRC cases and 118 controls free of colorectal neoplasms. The validation set included 41 CRC cases, 106 participants with advanced adenomas and 107 controls free of colorectal neoplasms, all of whom were recruited in the screening setting.

Table 1 provides the distribution of characteristics of the study population of the discovery set and the validation set. In both sets, CRC cases were on average a few years older than controls free of neoplasm and advanced adenomas. In

Table 1 Characteristics of the study population in the discovery set and validation set

| Group | Discovery set | | Validation set | | |
|---------------------------|---------------|----------------|----------------|-------------------------|----------------|
| | CRC (N, %) | Control (N, %) | CRC (N, %) | Advanced adenoma (N, %) | Control (N, %) |
| Total | 226 | 118 | 41 | 106 | 107 |
| Age (years) | | | | | |
| <60 | 50 (23.1) | 49 (42.6) | 7 (19.4) | 37 (36.3) | 46 (43.8) |
| 60–64 | 32 (14.8) | 29 (25.2) | 11 (30.6) | 31 (30.4) | 22 (21.0) |
| 65–69 | 30 (13.9) | 17 (14.8) | 5 (13.9) | 10 (9.8) | 17 (16.2) |
| ≥70 | 104 (48.1) | 20 (17.4) | 13 (36.1) | 24 (23.5) | 20 (19.0) |
| Mean ± SD | 67.8±12.0 | 62.0±7.3 | 66.8±6.9 | 63.5±7.1 | 62.2±6.3 |
| Sex | | | | | |
| Male | 129 (57.1) | 52 (44.1) | 29 (70.7) | 56 (52.8) | 44 (41.1) |
| Female | 97 (42.9) | 66 (55.9) | 12 (29.3) | 50 (47.2) | 62 (57.9) |
| UICC TNM tumor stage | | | | | |
| I | 67 (29.6) | – | 14 (34.1) | – | – |
| II | 64 (28.3) | – | 3 (7.3) | – | – |
| III | 67 (29.6) | – | 21 (51.2) | – | – |
| IV | 28 (12.4) | – | 3 (7.3) | – | – |
| CRC location | | | | | |
| Colon | 139 (61.5) | – | 23 (56.1) | – | – |
| Rectum | 87 (38.5) | – | 17 (41.5) | – | – |
| Unknown | – | – | 1 (2.4) | – | – |
| Advanced adenoma subclass | | | | | |
| HGD | – | – | – | 14 (13.2) | – |
| Villous | – | – | – | 56 (34.0) | – |
| Adenoma ≥ 1 cm | – | – | – | 36 (52.8) | – |

Abbreviations: CRC, colorectal cancer; HGD, high-grade dysplasia; UICC TNM, Union for International Cancer Control tumor node metastasis.

addition, the proportion of men was somewhat higher in the CRC group and in the advanced adenoma group than in the control group. Approximately half of the CRC cases were diagnosed in early (I or II) stages. A slightly higher proportion of CRCs was diagnosed at early stage (stage I/II) for the discovery set than for the validation set (57.9% vs 41.4%). More cancer patients had their tumor located in the colon than in the rectum.

Results of univariate analysis comparing the blood expression differences of the 92 individual protein biomarkers between CRC patients, advanced adenoma patients and controls free of neoplasm are summarized in [Tables S2 and S3](#). Overall, 39 proteins showed statistically significant different expression levels between CRC cases and controls free of neoplasm in the discovery set (adjusted *p*-values <0.05). A total of 12 of them were successfully replicated in the validation set even though included a much lower number of CRC cases (*n*=41). The respective results are listed in Table 2. All 12 proteins showed statistically significant higher expression levels in CRC cases than in controls. Two of them, i.e., GDF-15 and AREG, individually predicted the presence of CRC with an AUC>0.70.

We used the Lasso logistic regression models to construct a multi-marker prediction algorithm based on the 39 significant biomarkers identified in the discovery set. The following four proteins were selected in the algorithm: GDF-15, AREG, Fas antigen ligand (FasL) and Fms-related tyrosine kinase 3 ligand (Flt3L). Another prediction model combining these four proteins with TP53 autoantibody was further constructed. In the discovery set, the apparent AUCs of the four-marker algorithm and five-marker algorithm for discriminating CRC vs controls free of neoplasm were 0.87 (95% CI, 0.83–0.90) and 0.89 (95% CI, 0.85–0.92), respectively.

Table 3 and Figure 2 show the comparison of the two prediction algorithms in detecting CRC and its precursors in the validation set. The AUC of the four-protein panel for discriminating CRC vs controls free of neoplasm was 0.81 (95% CI, 0.73–0.88). Adding TP53 autoantibody to the four-protein panel conferred a modest improvement in terms of AUC (0.82, 95% CI, 0.74–0.90), but strong improvement could be observed at the left side of the ROC curve. When defining cutoffs yielding 90% specificity, the sensitivity of the four- and five-marker algorithm for detecting CRC was 53.6% (95% CI, 26.8%–70.7%) and 56.4% (95% CI,

Table 2 Diagnostic performance of 12 significant protein markers

| Protein marker | Discovery set | | Validation set | |
|----------------|---------------------------------------|----------|---------------------------------------|----------|
| | Fold change (CRC vs free of neoplasm) | p-value* | Fold change (CRC vs free of neoplasm) | p-value* |
| GDF-15 | 1.68 | <0.001 | 1.42 | <0.001 |
| AREG | 1.34 | <0.001 | 1.32 | <0.001 |
| TRAILR-2 | 1.23 | <0.001 | 1.28 | <0.001 |
| IL-6 | 1.56 | <0.001 | 1.48 | 0.014 |
| AM | 1.23 | <0.001 | 1.28 | 0.004 |
| HE4 | 1.24 | <0.001 | 1.19 | 0.014 |
| TNFR-2 | 1.21 | <0.001 | 1.16 | 0.014 |
| ILT3 | 1.21 | <0.001 | 1.14 | 0.043 |
| CEA | 1.29 | <0.001 | 1.86 | <0.001 |
| CXCL9 | 1.27 | 0.001 | 1.42 | 0.040 |
| TNFR-1 | 1.17 | 0.001 | 1.24 | 0.014 |
| HGF | 1.14 | 0.007 | 1.27 | 0.032 |

Note: *p-values were adjusted for multiple testing (Benjamini–Hochberg method).

Abbreviations: AM, adrenomedullin; AREG, amphiregulin; CEA, carcinoembryonic antigen; CRC, colorectal cancer; CXCL9, C-X-C motif chemokine 9; GDF-15, growth differentiation factor 15; HGF, hepatocyte growth factor; HE4, WAP four-disulfide core domain protein 2; IL-6, interleukin-6; ILT3, Immunoglobulin-like transcript 3; TNFR-1, tumor necrosis factor receptor-1; TNFR-2, tumor necrosis factor receptor-2; TRAILR-2, TNF-related apoptosis-inducing ligand receptor-2.

Table 3 Diagnostic performance of multi-marker panels for detecting CRC in the validation set

| Marker panel | AUC (95% CI) | Sensitivity (95% CI) | |
|--|------------------|----------------------|--------------------|
| | | At 90% specificity | At 80% specificity |
| CRC vs controls free of neoplasm | | | |
| Four-protein panel | 0.81 (0.73–0.88) | 53.6 (26.8–70.7) | 63.4 (48.8–82.9) |
| Four-protein panel + TP53 autoantibody | 0.82 (0.74–0.90) | 56.4 (38.4–71.8) | 66.7 (48.7–82.1) |
| Advanced adenomas vs controls free of neoplasm | | | |
| Four-protein panel | 0.58 (0.51–0.65) | 18.9 (8.5–27.4) | 23.6 (15.1–38.7) |
| Four-protein panel + TP53 autoantibody | 0.60 (0.52–0.69) | 22.0 (13.4–35.4) | 31.7 (18.3–45.1) |

Note: Four-protein panel includes GDF-15, AREG, Fas and Flt3L.

Abbreviations: AREG, amphiregulin; AUC, area under the curve; CRC, colorectal cancer; Flt3L, Fms-related tyrosine kinase 3 ligand; GDF-15, growth differentiation factor 15.

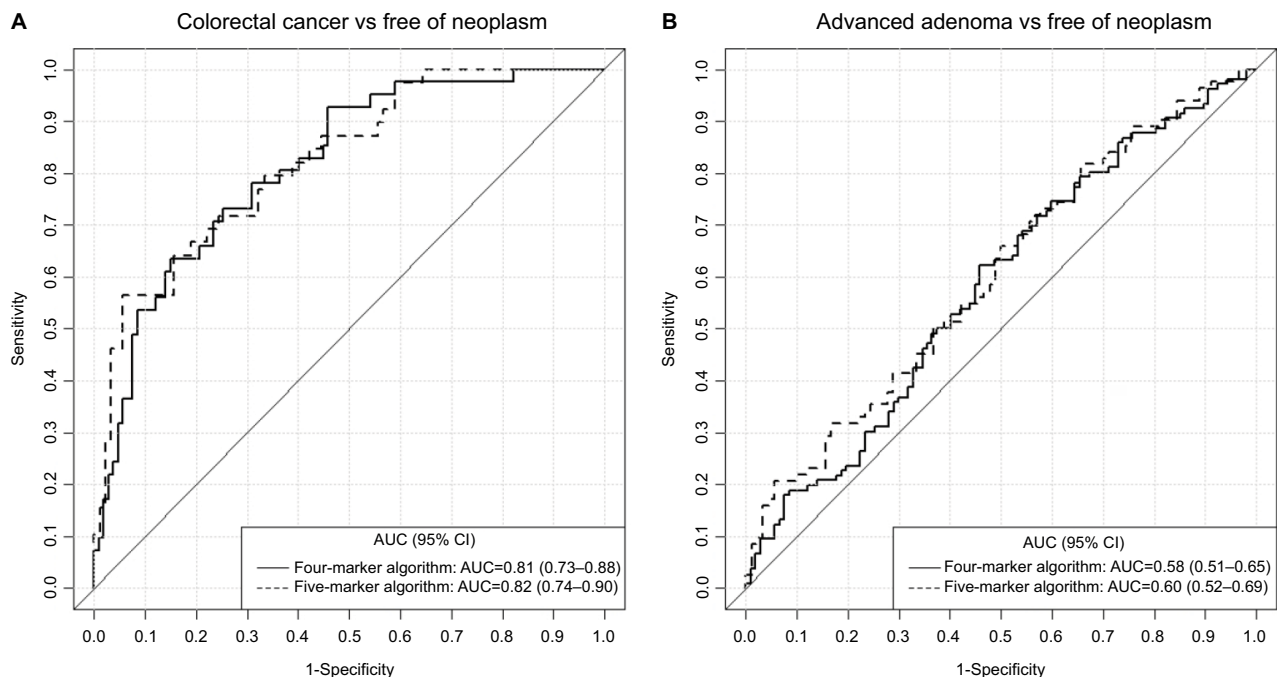


Figure 2 Comparison of ROC curves of the four- and five-marker panel for detecting: **(A)** CRC vs controls free of neoplasm; **(B)** advanced adenomas vs controls free of neoplasm.

Abbreviations: AUC, area under the curve; CRC, colorectal cancer; ROC, receiver operating characteristic.

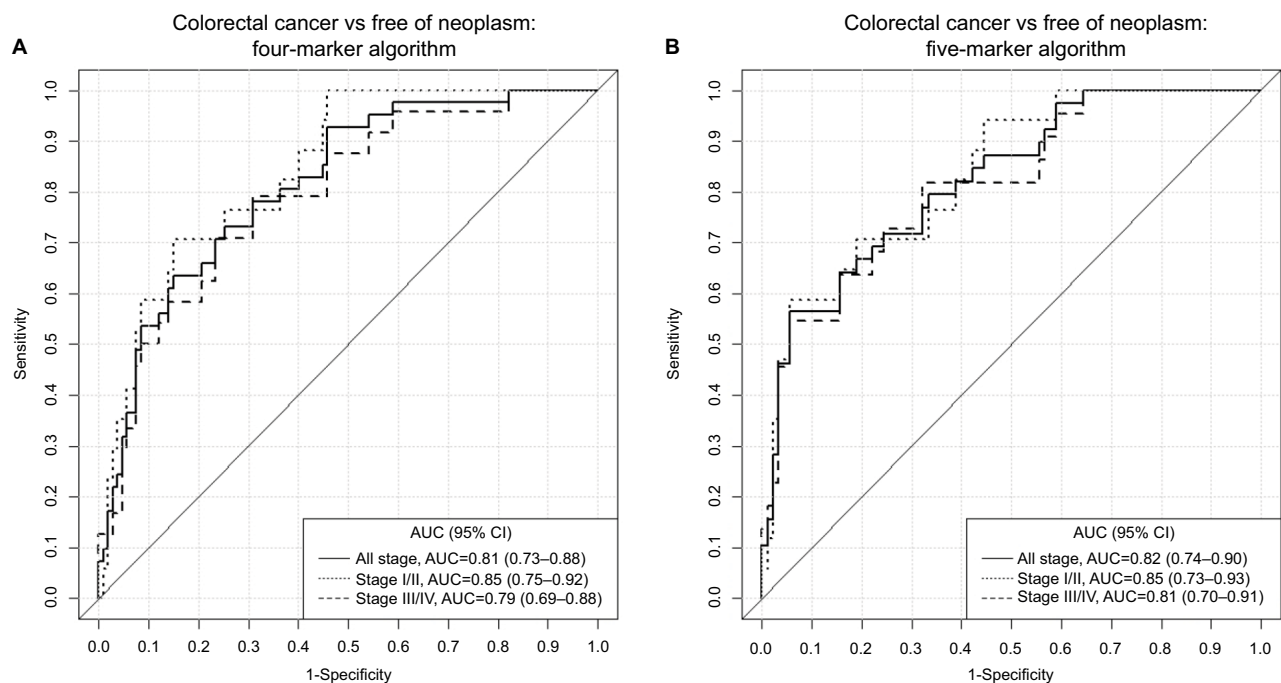


Figure 3 Comparison of ROC curves of (A) the four-marker panel and (B) five-marker panel for detecting early- and late-stage CRCs.

Abbreviations: AUC, area under the curve; CRC, colorectal cancer; ROC, receiver operating characteristic.

38.4%–71.8%), respectively, and at cutoffs yielding 80% specificity, the sensitivity of the four- and five-marker algorithm for detecting CRC was 63.4% (95% CI, 48.8%–82.9%) and 66.7% (95% CI, 48.7%–82.1%), respectively.

Both algorithms also showed modest diagnostic efficacy for detecting advanced adenomas, with AUCs of 0.58 (95% CI, 0.51–0.65) and 0.60 (95% CI, 0.52–0.69) for the four-protein algorithm and the five-marker algorithm, respectively. At the cutoffs yielding 90% specificity, the sensitivities of the four-protein algorithm and the five-protein algorithm for detecting advanced adenomas were 18.9% (95% CI, 8.5%–27.4%) and 22.0% (95% CI, 13.4%–35.4%), respectively.

Both prediction algorithms showed similar overall diagnostic performance for detecting early-stage CRC (TNM stage I/II) and late-stage CRC (TNM stage III/IV), as shown in Figure 3. For instance, the AUCs of the five-marker algorithm for detecting early- and late-stage CRC were 0.84 (95% CI, 0.73–0.92) and 0.81 (95% CI, 0.70–0.91), respectively. The differences were not statistically significant ($p=0.77$).

Sex- and age-specific results for the different outcomes are provided in [Tables S4 and S5](#). Both panels showed slightly higher AUCs for detecting CRC among women and in the younger age group (<65 years), but subgroup-specific CIs were wide and overlapping. Furthermore, adding age and sex to the five-marker panel in the regression models did not further improve the diagnostic performance for detecting CRC and its precursors ([Figure S1](#)).

Discussion

In this study, we evaluated the diagnostic performance of 92 plasma proteins and serum TP53 autoantibodies for detecting CRC and its precursors in a head-to-head manner using a large set of samples. A total of 12 protein biomarkers showed significantly higher expression levels in CRC patients than in controls free of neoplasm in both the discovery set and validation set which was entirely derived from a true screening setting. Moreover, a five-marker panel including GDF-15, AREG, FasL, Flt3L and TP53 autoantibody was constructed and validated. In the validation set, the AUCs of the five-marker panel for detecting CRC and advanced adenomas were 0.82 (95% CI, 0.74–0.90) and 0.60 (95% CI, 0.74–0.90), respectively. Of note, the panel showed similar diagnostic performance for detecting early- and late-stage CRCs.

For the 12 identified protein markers which showed significantly higher plasma levels in CRC patients than controls free of neoplasm, previous research has demonstrated involvement in various mechanisms relating to carcinogenesis, such as inflammation, anti-apoptosis and angiogenesis. A description of the biological function of these markers is provided in the [Supplementary materials](#).

Of the five markers included in the final panel, GDF-15 and AREG exhibited very good diagnostic performance for detecting CRC, with AUCs higher than 0.70 even when used as single prediction markers. GDF-15 (also known as

macrophage inhibitory cytokine-1) is a divergent member of the human TGF- β superfamily and a mediator of systemic inflammatory response²⁴ and has been reported to be related to various types of cancer.^{24–27} AREG belongs to the *EGF* family, which has been suggested to have pro-neoplastic effects in tissues of a wide variety of organs, such as colon, lung, and stomach, through proinflammatory mechanisms.^{28,29} Our results regarding GDF-15 are in line with previous studies.^{30,31} For instance, Mehta et al³⁰ reported that high levels of GDF-15 in blood were associated with increased risk of CRC in a nested case–control study including 618 incident CRC patients and 950 matched controls. Regarding AREG, previous research mostly focused on its role as a predictive marker in selecting patients for specific, targeted therapy.^{32,33} However, the current analysis along with our previous research¹³ also demonstrated that AREG carries excellent diagnostic potential for CRC detection and might be a good candidate to be included in multi-marker panels in the future.

The vast majority of studies evaluating blood-based biomarkers for CRC screening have been exclusively conducted in clinical settings. Some of these studies have reported higher levels of sensitivity and specificity,^{34,35} which typically could not be confirmed in rigorous validation in screening settings, however. In clinical settings, cases are typically symptomatic and have undergone a variety of diagnostic procedures (such as colonoscopy) leading to the diagnosis. Some patients may even have had initial therapeutic intervention before collection of blood samples. Controls often consist of or include patients with other diseases, and there may be differences in sample collection and processing procedures between cases and controls. All of these factors can influence apparent diagnostic performance and easily lead to false-positive findings.³⁶ We therefore paid utmost attention to avoid such bias by a rigorous study design in which independent validation of biomarkers identified in the discovery set was performed in a validation set that exclusively relied on participants recruited prior to screening colonoscopy in a true screening setting.

Advanced adenoma is the most important precursor of CRC, which a substantial risk of development into CRC in the long run.^{37–40} Early detection and removal of these precancerous lesions could therefore reduce the risk of CRC occurrence. To date, it is still a major challenge to detect advanced adenomas using blood-based tests, and most studies found very poor diagnostic performance for this outcome. Although some candidates, such as miRNA-135b³⁵ and a panel of BAG4, IL6ST and CD44,⁴¹ were reported in some studies to present good sensitivity for detecting advanced

adenomas, these findings were either derived from studies having limited sample size or using clinically identified cases, thus requiring further independently validation in larger screening populations. In our study, the five-marker panel presented limited diagnostic efficacy in detecting advanced adenomas, with an AUC of 0.60 (95% CI, 0.52–0.69). Even though the sensitivity for detecting advanced adenomas was higher than reported for other blood tests, major efforts should be made to identify blood-based tests with higher diagnostic accuracy for detecting advanced adenomas in addition to CRCs.

It should also be noted that the diagnostic performance of our proposed five-marker panel was still inferior to diagnostic performance of the most widely used stool test – fecal immunochemical test (FIT),^{42–44} and also a multi-target stool DNA test – Cologuard™, combining FIT with four stool DNA markers.^{45,46} However, our panel exhibited comparable diagnostic performance compared to the plasma methylated Septin 9, the only US FDA approved blood-based test for CRC screening. The sensitivity of methylated Septin 9 for detecting CRC and advanced adenomas were reported to be 48.2% and 11.2%, respectively, at a specificity of 91.5%.¹⁰ When adjusting the cutoff yielding the identical specificity as reported by Church et al,¹⁰ the sensitivities of the five-marker panel for detecting CRC and advanced adenomas were 56.4% and 20.7%, respectively (not reported in the “Results” section), suggesting slightly better diagnostic performance. Nevertheless, identification of supplementary markers that could further enhance the diagnostic performance of our algorithm to levels competitive with the best available stool tests in further research would be highly desirable.

Specific strengths and limitations deserve careful consideration when interpreting our results. Strengths include that we adopted a two-step approach, with biomarker discovery and subsequent validation in an independent sample set. Of note, the validation set consisted of prediagnostic samples from a large cohort of participants attending screening colonoscopy, therefore representing the target population for CRC screening. Moreover, both CRC and its precursors were included in the validation set, therefore rendering a thorough overview of the diagnostic potential of all examined biomarkers and the multi-marker panels. In addition, a large number of markers were tested simultaneously using state-of-the-art techniques, making a direct comparison of the diagnostic performance of all tested markers possible.

Limitations of our study include the small sample size of CRC included in the validation set, despite the very

large screening population recruited, reflecting the very low prevalence of CRC in a true screening population. Furthermore, to what extent the identified markers are CRC specific was not addressed in our study and should be assessed in further research. Therefore, before translation of these findings to clinical application, further assessment including patients with other diseases would be necessary. Finally, our analysis was based on the analysis of a single blood sample measurement by a multiplex assay per participant only. Given relatively high coefficients of variation (CV) reported for the protein measurements in the multiplex assay (average intra-CV was 5% and average inter-CV was 20% for our analysis), there seems to be potential to improve diagnostic performance by repeat measurements and by using other laboratory techniques (such as enzyme-linked immunosorbent assay) which should be explored in further research.

Conclusion

We identified several promising individual protein markers that carry diagnostic potential for the early detection of CRC. We also showed that a five-marker panel including GDF-15, AREG, FasL, Flt3L and TP53 autoantibody exhibited good diagnostic performance for detecting CRC and advanced adenomas. Although not competitive in diagnostic performance with the best-established stool-based CRC screening markers, the identified biomarkers could contribute to the development of a powerful blood-based test for CRC screening in the future.

Acknowledgments

The BLITZ study was partly funded by grants from the German Research Council (DFG, grant No. BR1704/16-1). The autoantibody measurements were supported by iMed – the Helmholtz Initiative on Personalized Medicine.

Disclosure

Hermann Brenner and Hongda Chen have applied for a patent “Biomarker panel for diagnosing cancer.” The other authors report no conflicts of interest in this work.

References

- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87–108.
- Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. *Gut*. 2015;64(5):784–790.
- Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ*. 2014;348:g2467.
- Zauber AG, Winawer SJ, O’Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012;366(8):687–696.
- Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol*. 2008;103(6):1541–1549.
- Bretthauer M, Kaminski MF, Loberg M, et al. Population-based colonoscopy screening for colorectal cancer: a randomized clinical trial. *JAMA Intern Med*. 2016;176(7):894–902.
- Klabunde C, Blom J, Bulliard JL, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. *J Med Screen*. 2015;22(3):119–126.
- Adler A, Geiger S, Keil A, et al. Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany. *BMC Gastroenterol*. 2014;14:183.
- US Food and Drug Administration [webpage on the Internet]. *Premarket Approval (PMA) for Epi proColon*. US Food and Drug Administration; 2017. Available from: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMA/pma.cfm?id=P130001>. Accessed September 15, 2017.
- Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63(2):317–325.
- Chen H, Werner S, Tao S, Zornig I, Brenner H. Blood autoantibodies against tumor-associated antigens as biomarkers in early detection of colorectal cancer. *Cancer Lett*. 2014;346(2):178–187.
- Luo X, Burwinkel B, Tao S, Brenner H. MicroRNA signatures: novel biomarker for colorectal cancer? *Cancer Epidemiol Biomarkers Prev*. 2011;20(7):1272–1286.
- Chen H, Zucknick M, Werner S, Knebel P, Brenner H. Head-to-head comparison and evaluation of 92 plasma protein biomarkers for early detection of colorectal cancer in a true screening setting. *Clin Cancer Res*. 2015;21(14):3318–3326.
- Brenner H, Tao S, Haug U. Low-dose aspirin use and performance of immunochemical fecal occult blood tests. *JAMA*. 2010;304(22):2513–2520.
- Hundt S, Haug U, Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. *Ann Intern Med*. 2009;150(3):162–169.
- Pox CP, Altenhofen L, Brenner H, Theilmeyer A, Von Stillfried D, Schmiegel W. Efficacy of a nationwide screening colonoscopy program for colorectal cancer. *Gastroenterology*. 2012;142(7):1460.e2–1467.e2.
- Brenner H, Altenhofen L, Kretschmann J, et al. Trends in adenoma detection rates during the first 10 years of the German Screening Colonoscopy Program. *Gastroenterology*. 2015;149(2):356.e1–366.e1.
- Brenner H, Altenhofen L, Tao S. Matching of controls may lead to biased estimates of specificity in the evaluation of cancer screening tests. *J Clin Epidemiol*. 2013;66(2):202–208.
- Tao S, Haug U, Kuhn K, Brenner H. Comparison and combination of blood-based inflammatory markers with faecal occult blood tests for non-invasive colorectal cancer screening. *Br J Cancer*. 2012;106(8):1424–1430.
- Chen H, Werner S, Butt J, et al. Prospective evaluation of 64 serum autoantibodies as biomarkers for early detection of colorectal cancer in a true screening setting. *Oncotarget*. 2016;7(13):16420–16432.
- Zornig I, Halama N, Lorenzo Bermejo J, et al. Prognostic significance of spontaneous antibody responses against tumor-associated antigens in malignant melanoma patients. *Int J Cancer*. 2015;136(1):138–151.
- Waterboer T, Sehr P, Michael KM, et al. Multiplex human papillomavirus serology based on in situ-purified glutathione s-transferase fusion proteins. *Clin Chem*. 2005;51(10):1845–1853.
- R Core Team [homepage on the Internet]. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014. Available from: <http://www.R-project.org/>. Accessed September 15, 2017.
- Bootcov MR, Bauskin AR, Valenzuela SM, et al. MIC-1, a novel macrophage inhibitory cytokine, is a divergent member of the TGF-beta superfamily. *Proc Natl Acad Sci U S A*. 1997;94(21):11514–11519.

25. Song M, Mehta RS, Wu K, et al. Plasma inflammatory markers and risk of advanced colorectal adenoma in women. *Cancer Prev Res (Phila)*. 2016;9(1):27–34.
26. Brown DA, Hance KW, Rogers CJ, et al. Serum macrophage inhibitory cytokine-1 (MIC-1/GDF15): a potential screening tool for the prevention of colon cancer? *Cancer Epidemiol Biomarkers Prev*. 2012;21(2):337–346.
27. Staff AC, Trovik J, Eriksson AG, et al. Elevated plasma growth differentiation factor-15 correlates with lymph node metastases and poor survival in endometrial cancer. *Clin Cancer Res*. 2011;17(14):4825–4833.
28. Berasain C, Avila MA. Amphiregulin. *Semin Cell Dev Biol*. 2014;28:31–41.
29. Busser B, Sancey L, Brambilla E, Coll JL, Hurbain A. The multiple roles of amphiregulin in human cancer. *Biochim Biophys Acta*. 2011;1816(2):119–131.
30. Mehta RS, Song M, Bezawada N, et al. A prospective study of macrophage inhibitory cytokine-1 (MIC-1/GDF15) and risk of colorectal cancer. *J Natl Cancer Inst*. 2014;106(4):dju016.
31. Wallin U, Glimelius B, Jirstrom K, et al. Growth differentiation factor 15: a prognostic marker for recurrence in colorectal cancer. *Br J Cancer*. 2011;104(10):1619–1627.
32. Seligmann JF, Elliott F, Richman SD, et al. Combined epiregulin and amphiregulin expression levels as a predictive biomarker for panitumumab therapy benefit or lack of benefit in patients with RAS wild-type advanced colorectal cancer. *JAMA Oncol*. Epub 2016 Feb 11.
33. Yonesaka K, Takegawa N, Satoh T, et al. Combined analysis of plasma amphiregulin and heregulin predicts response to cetuximab in metastatic colorectal cancer. *PLoS One*. 2015;10(11):e0143132.
34. Wong SH, Kwong TN, Chow TC, et al. Quantitation of faecal Fusobacterium improves faecal immunochemical test in detecting advanced colorectal neoplasia. *Gut*. 2017;66(8):1441–1448.
35. Wu CW, Ng SC, Dong Y, et al. Identification of microRNA-135b in stool as a potential noninvasive biomarker for colorectal cancer and adenoma. *Clin Cancer Res*. 2014;20(11):2994–3002.
36. Chen H, Knebel P, Brenner H. Empirical evaluation demonstrated importance of validating biomarkers for early detection of cancer in screening settings to limit the number of false-positive findings. *J Clin Epidemiol*. 2016;75:108–114.
37. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet*. 2014;383(9927):1490–1502.
38. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med*. 1988;319(9):525–532.
39. Morson B. The polyp-cancer sequence in the large bowel. *Proc R Soc Med*. 1974;67(6 pt 1):451–457.
40. Brenner H, Altenhofen L, Stock C, Hoffmeister M. Natural history of colorectal adenomas: birth cohort analysis among 3.6 million participants of screening colonoscopy. *Cancer Epidemiol Biomarkers Prev*. 2013;22(6):1043–1051.
41. Rho JH, Ladd JJ, Li CI, et al. Protein and glycomic plasma markers for early detection of adenoma and colon cancer. *Gut*. Epub 2016 Nov 7.
42. Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. *Ann Intern Med*. 2014;160(3):171.
43. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut*. 2015;64(8):1327–1337.
44. Robertson DJ, Lee JK, Boland CR, et al. Recommendations on fecal immunochemical testing to screen for colorectal neoplasia: a consensus statement by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology*. 2017;152(5):1217.e3–1237.e3.
45. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med*. 2014;370(14):1287–1297.
46. Dickinson BT, Kisiel J, Ahlquist DA, Grady WM. Molecular markers for colorectal cancer screening. *Gut*. 2015;64(9):1485–1494.

Clinical Epidemiology

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventive initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>

systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Dovepress