


RESEARCH ARTICLE

Open Access



A fecal-based test for the detection of advanced adenoma and colorectal cancer: a case-control and screening cohort study

Lian-Jing Cao^{1,2†}, Xiao-Lin Peng^{3†}, Wen-Qiong Xue^{1†}, Rong Zhang^{4†}, Jiang-Bo Zhang¹, Ting Zhou^{1,5}, Zi-Yi Wu¹, Gai-Rui Li³, Tong-Min Wang¹, Yong-Qiao He¹, Da-Wei Yang⁶, Ying Liao¹, Xia-Ting Tong⁶, Fang Wang⁷, Ke-Xin Chen⁸, Shi-Hong Zhang⁹, Li-Qing Zhu³, Pei-Rong Ding¹⁰ and Wei-Hua Jia^{1,5,6*} 

Abstract

Background: Colorectal cancer (CRC) is the leading cause of cancer death worldwide. Screening is a confirmed way to reduce the incidence and mortality rates of CRC. This study aimed to identify a fecal-based, noninvasive, and accurate method for detection of colorectal cancer (CRC) and advanced adenoma (AA).

Methods: Through detection in tissue ($n = 96$) and fecal samples ($n = 88$) and tested in an independent group of fecal samples ($n = 294$), the methylated DNA marker ITGA4 and bacterial markers *Fusobacterium nucleatum* (*Fn*) and *Peptostreptococcus anaerobius* (*Pa*) were identified from the candidate biomarkers for CRC and AA detection. A prediction score (pd-score) was constructed using the selected markers and fecal immunochemical test (FIT) for distinguishing AA and CRC from healthy subjects by logistic regression method. The diagnostic performance of the pd-score was compared with FIT and validated in the external validation cohort ($n = 117$) and in a large CRC screening cohort.

Results: The pd-score accurately identified AA and CRC from healthy subjects with an area under the curve (AUC) of 0.958, at a specificity of 91.37%; the pd-score showed sensitivities of 95.38% for CRC and 70.83% for AA, respectively. In the external validation cohort, the sensitivities of the pd-score for CRC and AA detection were 94.03% and 80.00%, respectively. When applied in screening, the pd-score identified 100% (11/11) of CRC and 70.83% (17/24) of AA in participants with both colonoscopy results and qualified fecal samples, showing an improvement by 41.19% compared to FIT.

Conclusions: The current study developed a noninvasive and well-validated approach for AA and CRC detection, which could be applied widely as a diagnostic and screening test.

Keywords: Colorectal cancer, Advanced adenoma, Noninvasive test, Fecal biomarkers, Screening

* Correspondence: jjawh@sysucc.org.cn

[†]Lian-Jing Cao, Xiao-Lin Peng, Wen-Qiong Xue and Rong Zhang contributed equally to this work.

¹State Key Laboratory of Oncology in South China Guangzhou, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, People's Republic of China

⁵Biobank of Sun Yat-sen University Cancer Center, Guangzhou, People's Republic of China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Colorectal cancer (CRC) is the third most commonly diagnosed and the second most fatal cancer, contributing to approximately 10.2% of the annual cancer incidence and 9.2% of the cancer-related mortalities [1, 2]. With the development of socioeconomics and changes in dietary patterns, the incidence and mortality rates of CRC have increased in China in recent years [3].

The underlying neoplastic process from an aberrant crypt to a precursor lesion and eventually to CRC takes 10 to 15 years, providing an optimal window phase for CRC screening [4]. Numerous studies support the fact that screening contributes to the discovery and removal of precursor lesions, represented by adenomas, and early-stage CRC, which could reduce the incidence and mortality rates of CRC. In fact, approximately 63% of CRC deaths could be attributed to the lack of regular screening [5]. Data in the USA show an increase in CRC screening rate from 38% in 2000 to 66% in 2018, which correspondingly resulted in a substantial decrease in CRC incidence and CRC-related mortality [6].

Colonoscopy, a representative structural-based examination, is limited in large-scale population CRC screening due to its high invasiveness and time consumption. FIT is now the most commonly applied noninvasive test due to its easy operation and ability to reduce CRC-related mortality [7]. However, the sensitivity of FIT for early-stage CRC detection, especially for adenomas, is rather unsatisfactory. A meta-analysis demonstrated a sensitivity of 79% (95% CI 0.69 to 0.86) for CRC in average-risk populations [8], while the sensitivity was only 6% to 56% for AA [9]. Thus, in the widely accepted 2-step screening scenario (a simple test finding out those need colonoscopy), the relatively low sensitivity of FIT leads to a relatively high false-negative rate.

The liquid biopsy provides a noninvasive route of sample collection for analysis of tumor-derived DNA, RNA, miRNA, and proteins. A number of studies suggest that analysis of tumor-derived DNA, RNA, miRNA, or proteins can provide relevant information for CRC detection and some tests have been commercially available, including the ColonSentry™ messenger RNA (mRNA) expression panel [10], the Colox® 29-gene panel [11], the CELTiC panel [12], and the *SEPT9* methylated DNA test [13]. However, besides variations regarding the blood collection processing and sample storage, the sensitivity of the abovementioned commercialized kits for identifying CRC from healthy is still have room to improvement.

The stool DNA (sDNA) test, which detects alterations in DNA in tumor cells that slough into the stool, is a new approach for CRC diagnosis and screening. It is a desirable CRC screening method due to its noninvasive, highly sensitive, and user-friendly properties. In 2014,

the first sDNA detection kit, Cologuard™ (Exact Science, Madison WI), was approved by the US Food and Drug Administration (FDA) for clinical use [14]. To date, the sDNA test has been recommended as a CRC screening method in the NCCN [15] and Chinese guidelines [16]. For countries with large populations, such as China, it is reasonable to adopt a specific biomarker-based screening kit as a frontline test to select suspected CRC patients and then carry out colonoscopy.

A wide variety of genetic and molecular changes mediate colorectal carcinogenesis, including defective DNA repair, chromosome instability, microsatellite instability, DNA methylation, and microbiota [17]. Many studies have shown compelling evidence that epigenetic markers are ideal diagnostic biomarkers, as they appear quite early in disease pathogenesis [18]. DNA methylation is one of the most ubiquitous epigenetic changes in carcinogenesis [19]. Hypermethylation occurring in CpG islands plays a key pathophysiological role in the initiation and progression of CRC [20]. Scientists have reported many highly sensitive methylated DNA markers for CRC detection and some of them have been commercialized [21–23], including the FDA approved Cologuard™.

Function and application of gut microbiota have been widely explored in recent years. Studies on the pathogenic mechanisms of bacteria have revealed that several pathogenic microbiota contribute to colorectal carcinogenesis. Consistently, the composition of the gut microbiota is different between CRC patients and healthy individuals; some potential protective taxa are decreased, while other procarcinogenic taxa are increased in CRC [24–26]. Although geographically variant, several bacterial species are reproducible and invariably enriched in CRC [27, 28], which pinpoints the potential diagnostic value of detecting a core set of bacteria.

This study aimed to identify a fecal-based, effective, and accurate method for CRC and AA diagnosis and screening. Here, we combined the highly sensitive sDNA test and the highly specific FIT and constructed a prediction model, the pd-score. Moreover, to detect more CRC and AA patients, both CRC-specific methylation markers and pathogenic bacterial markers were included in the sDNA test. The prediction model was well-validated in an external validation cohort and applied in a CRC screening cohort.

Methods

Study design

To obtain a reliable CRC diagnosis and screening method, our study consisted of the following two main phases: (1) Model construction phase: the candidate methylated DNA and bacterial markers were first

quantified in tissue and fecal samples collected from CRC patients and healthy subjects; the selected markers were then evaluated in an independent group of fecal samples collected from CRC, AA, and healthy subjects; then, the pd-score was constructed for CRC and AA

diagnosis using a logistic regression model of the fecal samples mentioned above. (2) Validation phase: The pd-score was validated in the external validation cohort in which fecal samples were collected from three institutions and in a large CRC screening cohort (Fig. 1A, B).

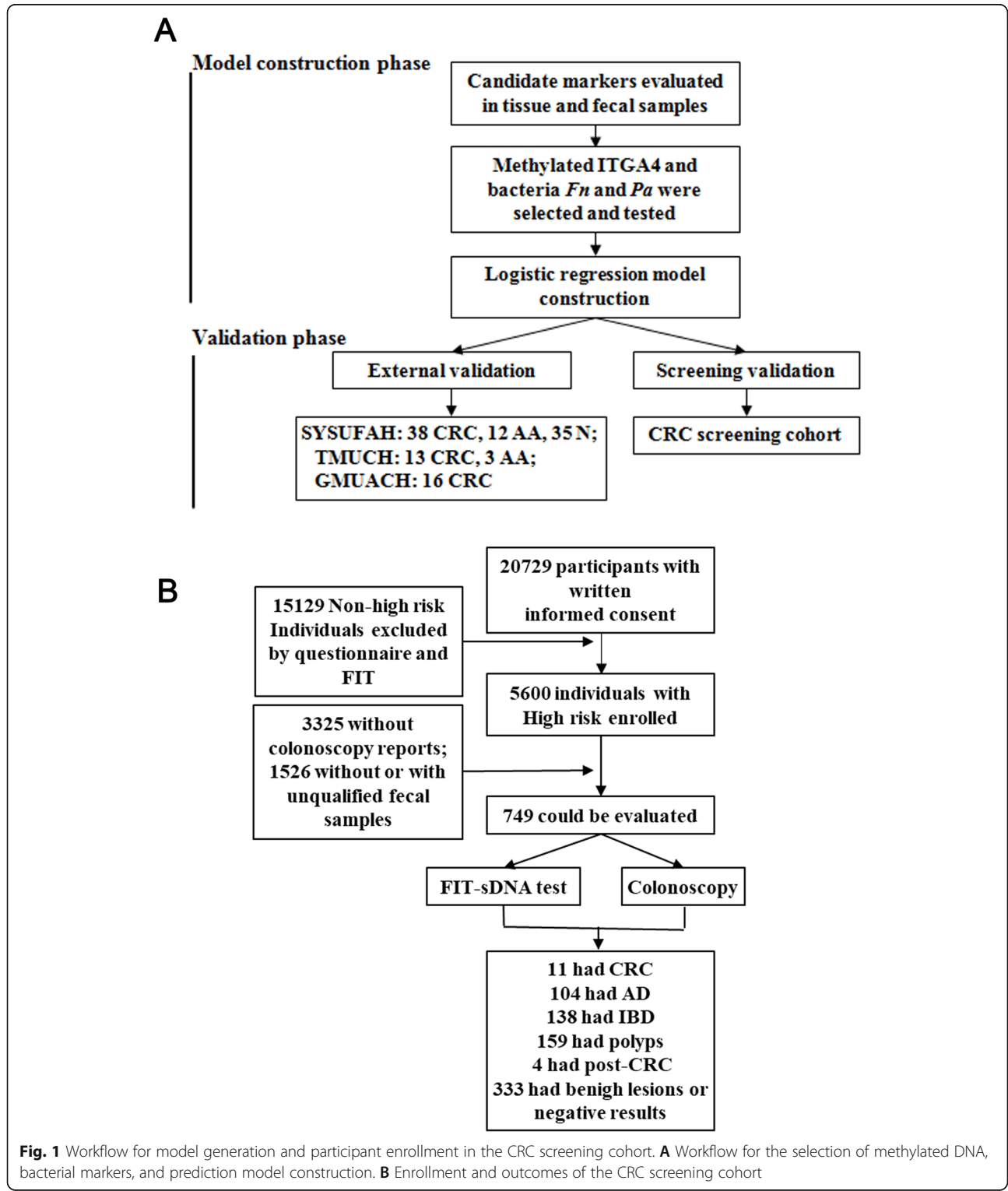


Fig. 1 Workflow for model generation and participant enrollment in the CRC screening cohort. **A** Workflow for the selection of methylated DNA, bacterial markers, and prediction model construction. **B** Enrollment and outcomes of the CRC screening cohort

Study population

Paired cancer and adjacent normal mucosa tissues obtained from 48 stage I–II CRC patients were collected from Sun Yat-sen University Cancer Center (SYSUCC). Fecal samples used in the model construction phase included 195 CRC and 48 AA patients and 139 healthy subjects collected from SYSUCC, which was named the SYSUCC cohort. Fecal samples of the external validation cohort were collected from subjects from three institutions, including 38 CRC and 12 AA patients and 35 healthy subjects from the SYSUFAH, 16 CRC patients from the Affiliated Cancer Hospital and Institute of Guangzhou Medical University (GMUACH), and 13 CRC and 3 AA patients from Tianjin Cancer Institute & Hospital (TMUCH). The demographics of the SYSUCC cohort and the external validation cohort are summarized in Additional file 2: Table S1 and Table S2.

Fecal samples were collected either 1 week before or 1 month after colonoscopy to allow recovery of the gut microbiome [29]. Hospital-based subjects were recruited according to the following criteria: (1) pathologically diagnosed with CRC or AA, (2) the pathological type was adenocarcinoma or adenoma, and (3) no prior history of antitumor treatment. The exclusion criteria were as follows: (1) use of antibiotics within the past 3 months, (2) history of or the presence of any other cancer, (3) a formless fecal sample, (4) inflammatory disease of the gut, and (5) any invasive medical intervention (including resection of adenoma or polyp during colonoscopy, while not only colonoscopy) within the past 3 months. Prior consent from all participants and approval from the Research Ethics Committee of the SYSUCC were obtained for experimentation with human samples.

The stages of the CRC patients were determined according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system. In our study, the adenoma was diagnosed by experienced pathologist through review of histological features. The adenomas in this study include both advanced adenoma and non-advanced adenoma. Advanced adenomas included the adenomas ≥ 1 cm along the greatest dimension, or with $\geq 25\%$ villous histological features, or high-grade dysplasia. Other histologically diagnosed adenomas not meeting these standards were defined as non-advanced adenoma [14]. Proximal CRCs were defined as tumors located in the cecum, ascending colon or hepatic flexure, and distal CRCs were defined as those located in the splenic flexure, descending colon, sigmoid colon, or rectum.

CRC screening cohort

The CRC screening cohort originated from a subset of the Cancer Screening Program in Urban China [30] and was collected from Nanshan District, Shenzhen city,

Guangdong Province, China. Participants enrolled in this study were people aged 45–75 years. All participants were first invited to take a risk assessment using an established Clinical Cancer Risk Score System [30] and FIT. Those who were assessed as positive by the Clinical Cancer Risk Score System or FIT were considered to be at high-risk for CRC and were then recommended to undergo colonoscopy and to provide fecal samples within 90 days after the risk assessment. A total of 20729 participants were enrolled in this study from May 2017 to December 2019. A total of 5600 participants were considered to be at high-risk for CRC, and finally, 749 participants with colonoscopy results and qualified fecal samples could be evaluated. This cohort contained 11 CRC, 104 adenoma (AD), 159 colorectal polyps, 138 inflammatory bowel disease (IBD), 4 postoperative CRC (post-CRC), and 333 other benign lesions or negative results. The demographics of the screening cohort are summarized in Additional file 2: Table S3. This project was approved by the Ethics Committee of the Shenzhen Nanshan Center for Chronic Disease Control, and prior consent was obtained from all participants.

Fecal sample collection and storage

Fecal samples were collected by two different methods: immediate freezing for hospital-based samples and mixing in fecal preservative buffer (FPB) for the screening samples. Samples from patients and healthy subjects at the hospital were allocated and frozen at -80°C within 4 h after defecation. People enrolled in the screening subject were asked to fill a 2-ml tube containing 1 ml FPB with fecal samples. Buffered samples were delivered to the laboratory and stored at -80°C within 24 h. There was no difference in human genome DNA integrity (Additional file 1: Figure S1A) or bacterial diversity (Additional file 1: Figure S1B–E) between the immediate freezing and buffered samples.

Marker selection

Seven methylated DNA and five bacterial markers were candidates based on literature review and available CRC detection kit using stool samples. For methylation markers, VIM, BMP3, NDRG4, and SDC2 came from the commercialized CRC detection kits [14, 31–33]. ITGA4 [34], MAL, and CNRIP1 [35] were included because they exhibited good separation of CRC from normal tissues.

For bacterial markers, *Fusobacterium nucleatum* (*Fn*), *Solobacteriummoorei* (*Sm*), *Pepetostreptococcusanaerobius* (*Pa*), and *Parvimonasmicra* (*Pm*) were included in this study because they were the most reproducible bacterial biomarkers for CRC detection across different datasets [27]. In addition, *Clostridium hathewayi* (*Ch*) was included as it was recently reported to be significantly

enriched in patients with colorectal adenoma compared with healthy subjects [36].

DNA extraction, quantitative real-time PCR of methylated DNA, and bacterial markers

Tissue DNA was isolated with the QIAamp DNA Mini Kit (Qiagen, Valencia, CA) according to the published protocol. Fecal samples (200 mg or 200 μ l) were used to isolate human and pathogen DNA with a QIAamp DNA Fecal Mini Kit (Qiagen, Valencia, CA). Briefly, to ensure fecal sample is thoroughly homogenized, we added 5 grinding beads to each stool sample after adding 1 ml InhibitEX buffer, and the samples were vortex continuously for 2 min. And the following operation was according to the manufacturer's protocol. Tissue- and fecal-derived DNA were chemically modified with sodium bisulfite using an EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA).

Quantitative real-time PCR amplification was adopted for the measurement of methylation status and the abundances of bacterial markers. For methylation markers, a 20- μ l reaction system of EpiTect MethyLight Mix (Qiagen, Valencia, CA) containing 5- μ l converted DNA, 250 nM of each primer, and 200 nM of each probe was applied; the thermal cycler parameters of the *LightCycler 480* instrument II were 95°C for 5 min and (95°C for 15 s and 60°C for 1 min) \times 45 cycles. Detection of bacterial abundances was performed with LightCycler 480 SYBR Green I Master Mix (Roche, Applied Science) in a 10- μ l reaction system containing 30–90 ng DNA and 250 nM of each primer; thermal cycler parameters were 95°C for 5 min and (95°C for 15 s and 60°C for 1 min) \times 45 cycles. Nucleotide sequences of the primers and probes are listed in Additional file 2: Table S4 and Table S5 [36–38].

For tissue samples, we use percent of methylated reference (PMR) to assess the methylation level of target genes. Plasmid containing the bisulfate-treated target gene region was constructed and the copy number of each plasmid was calculated according to the following formula: the number of copies of DNA template per μ l = (DNA concentration (ng/ μ l) \times Avogadro's number) / (length of template (bp) \times 10e9 \times 660). The plasmids were then serially diluted as standards for absolute quantification (10^1 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , and 10^7 copies per μ l). Bisulfite-treated CpGenome Universal Methylated DNA was used as positive control. The methylation level of each target gene was calculated as follows: methylation level = (target sample/ ACTB sample) / (target_{positive}/ ACTB_{positive}) \times 100%.

For fecal samples, the quantified CT value was used to directly interpret the level of each methylated gene. Samples without CT values of methylated genes were given a value of 45 to compare the methylation level

between different samples. Bacterial markers are presented as abundances and are relative units normalized to 16S rDNA using the $2^{-\Delta Ct}$ method ($\Delta Ct = Ct$ bacterial markers – Ct 16S rDNA and relative abundances = $2^{-\Delta Ct}$), which was performed as in a previously described protocol [39]. The relative abundance of each marker is shown as a log value of " $\times 10^8 + 1$ " [36].

Fecal immunochemical test

The FIT was performed using the Fecal Occult Blood Gold Gel Stripe (W.H.P.M. BIORESEARCH & TECHNOLOGY) according to the manufacturer's protocol. The minimum allowable detection of hemoglobin was 200 ng per milliliter. Samples were considered positive when color bands appeared simultaneously in the control line and the reaction line; those with only one color band in the control line were determined to be negative, while samples with no color band or only one color band that appeared in the reaction line were regarded as invalid and needed another test. The results are valid within 5 min. The laboratory staff performing the test was experienced and blinded to the colonoscopy results.

Quality control of the fecal samples

Human HCT116 DKO methylated and nonmethylated DNA (Zymo Research, Irvine CA) were used as positive and negative controls for the PCR system. Samples were considered valid if they satisfied all of the following requirements: (1) CT value of ACTB < 40, (2) CT value of 16S rDNA < 35, and (3) positive control line of the Fecal Occult Blood Gold Gel Stripe in the FIT test.

Statistical analysis

The Mann-Whitney *U* test was employed to calculate the differences in gene methylation level and bacterial abundance between the two groups of subjects. Kruskal-Wallis H test was performed to estimate the difference of ITGA4 methylation as well as the abundances *Fn* and *Pa* between healthy subjects, advanced adenoma and cancer. The receiver operating characteristic (ROC) curve was used to evaluate the diagnostic performance of markers or models in discriminating patients from healthy subjects. The best cutoff values of the ROC curve were determined by the maximal Youden index (sensitivity+specificity–1). Pairwise comparison of ROC curves was calculated with the method of DeLong et al. [40]. A logistic regression model was used to estimate the performance of marker combinations in discriminating patients from healthy subjects. The pd-score generated from the optimal logistic regression model was calculated as follows: logit (pd-score) = $\alpha + \beta_1 * ITGA4 + \beta_2 * Fn + \beta_3 * Pa + \beta_4 * FIT$, where α represents the intercept and β represents the regression coefficients of each marker. The Mann-Whitney *U* test was used to calculate

the differences in the pd-scores of the two groups of subjects. A one-sided McNemar paired-comparisons test was performed to observe the difference in sensitivity between the pd-score and FIT. The chi-square test was applied to assess the association of the detection rate with clinical covariates. Integrated discrimination improvement (IDI) was used to judge the improvement in the pd-score compared with FIT. All hypothesis tests, excluding the McNemar paired-comparisons test, were conducted in a two-sided manner, and a P value < 0.05 was considered to be statistically significant. All analyses were performed in R software, version 3.6.3.

Results

Identification of the methylated DNA ITGA4 and bacterial *Fn* and *Pa* as diagnostic markers for CRC

To identify better methylated DNA markers, we first detected the methylation level of the seven candidate markers in 48 CRC and adjacent normal tissues. The methylation levels of all seven candidates were significantly higher in CRC tissues than in adjacent normal tissues (all $P < 0.001$) (Fig. 2A). Methylated CNRIP1, ITGA4, and MAL showed areas under the ROC curve (AUCs) greater than 0.95 (AUC = 0.982 for CNRIP1, AUC = 0.974 for ITGA4 and AUC = 0.969 for MAL,

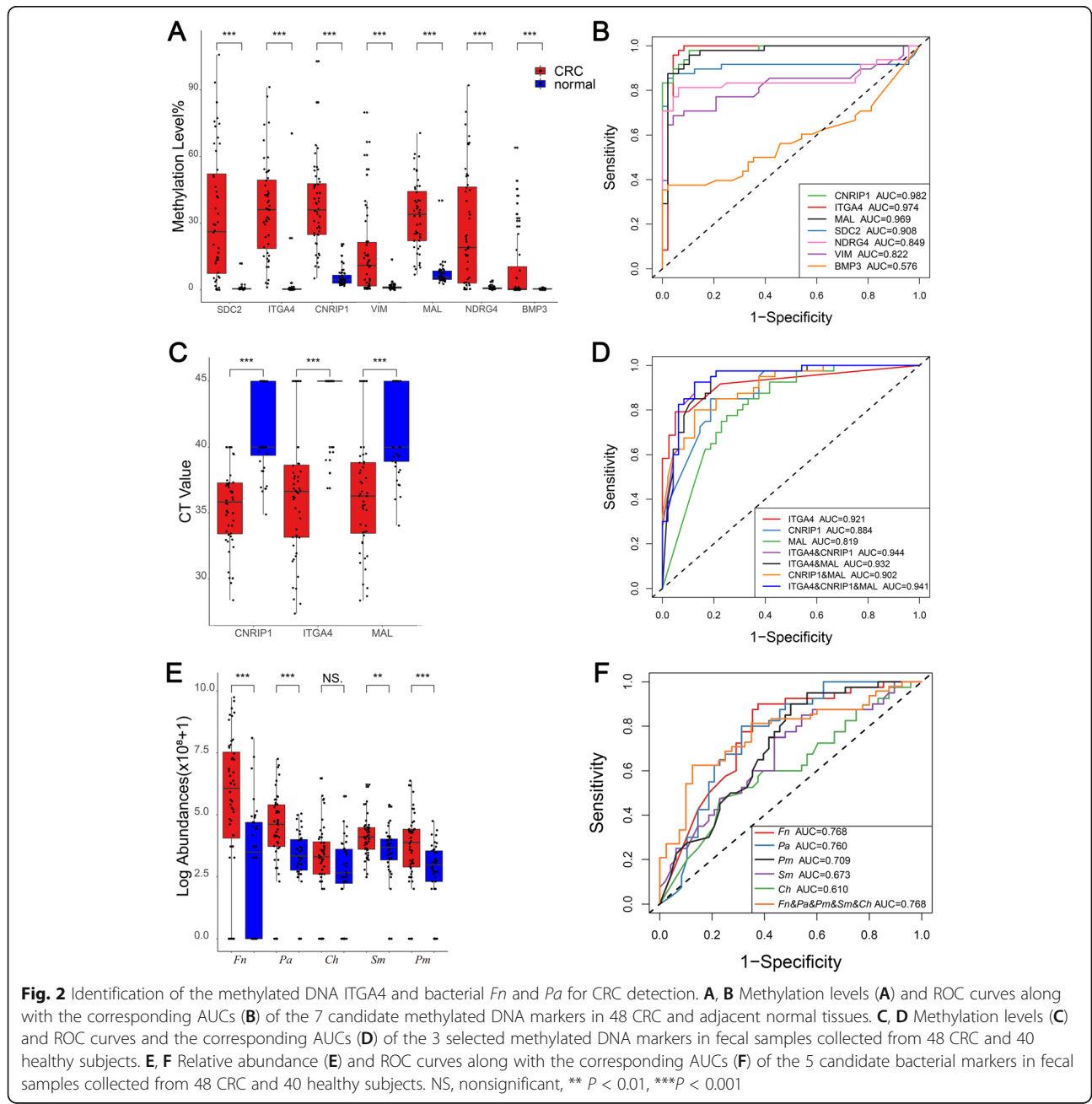


Fig. 2 Identification of the methylated DNA ITGA4 and bacterial *Fn* and *Pa* for CRC detection. **A, B** Methylation levels (**A**) and ROC curves along with the corresponding AUCs (**B**) of the 7 candidate methylated DNA markers in 48 CRC and adjacent normal tissues. **C, D** Methylation levels (**C**) and ROC curves and the corresponding AUCs (**D**) of the 3 selected methylated DNA markers in fecal samples collected from 48 CRC and 40 healthy subjects. **E, F** Relative abundance (**E**) and ROC curves along with the corresponding AUCs (**F**) of the 5 candidate bacterial markers in fecal samples collected from 48 CRC and 40 healthy subjects. NS, nonsignificant, ** $P < 0.01$, *** $P < 0.001$

respectively) and were selected for further testing (Fig. 2B).

To elucidate whether methylation markers that perform well in tissues could be reproduced in fecal samples, we detected methylated CNRIP1, ITGA4, and MAL in fecal samples collected from 48 CRC patients and 40 normal controls. All three markers showed significantly higher aberrant methylation levels in CRC patients than in healthy subjects (all $P < 0.001$) (Fig. 2C). ROC curve analysis demonstrated that ITGA4 had the highest diagnostic efficiency with an AUC of 0.921, but the combination of the three markers (CNRIP1 and ITGA4 and MAL) did not significantly enhance the diagnostic efficiency (AUC = 0.921 for ITGA4 vs. AUC = 0.941 for CNRIP1 and ITGA4 and MAL, $P = 0.163$) (Fig. 2D). Meanwhile, 31 out of 40 normal controls showed an absence of ITGA4 methylation, indicating a high specificity of methylated ITGA4 in the fecal samples (Fig. 2C).

We further quantitatively examined the abundances of the five candidate bacterial markers using the fecal samples mentioned above. Among the five candidates, except for *Ch*, the bacterial markers *Fn*, *Pa*, *Sm*, and *Pm* were all significantly enriched in CRC patients compared to healthy subjects ($P < 0.001$ for *Fn*, *Pa* and *Pm*; $P = 0.005$ for *Sm*) (Fig. 2E). ROC curve analysis showed that *Fn* and *Pa* performed better than *Sm* and *Pm* in discriminating CRC from healthy subjects, with AUCs of 0.768 and 0.760, respectively (Fig. 2F). The combination of the five markers (*Fn*, *Pa*, *Ch*, *Pm*, and *Sm*) yielded an AUC value of 0.768, which was not significantly enhanced compared with *Fn* or *Pa* alone.

Testing of the selected markers in an independent group

We further tested the diagnostic reliability of methylated ITGA4 and bacteria *Fn* and *Pa* in an independent group of fecal samples including 147 CRC, 48 AA, and 99 normal controls. Surprisingly, the levels of ITGA4, *Fn* and *Pa* were not only significantly higher in patients with CRC than in healthy subjects (all $P < 0.001$) but also significantly enhanced in patients with AA compared with healthy subjects (all $P < 0.001$) (Fig. 3A–C).

ROC curve analysis demonstrated that ITGA4, *Fn*, and *Pa* exhibited stable diagnostic performance for the detection of CRC, yielding AUCs of 0.894, 0.791, and 0.800, respectively (Fig. 3D). Furthermore, ITGA4, *Fn*, and *Pa* also performed well in discriminating AA patients from healthy subjects, with AUCs of 0.815, 0.703, and 0.669, respectively (Fig. 3E). These results verified the reliability of the diagnostic efficiency of ITGA4, *Fn*, and *Pa*.

Construction of the prediction model for AA and CRC diagnosis

As combinations of different methylation markers or bacterial markers are unable to significantly improve the

diagnostic efficiency, we then considered the possibility that combinations of different types of markers could achieve a better diagnostic performance. A prediction model was constructed using fecal samples from the SYSUCC cohort, including 195 CRC and 48 AA patients and 139 normal controls. The results showed that among all the different combinations, the logistic model constructed with ITGA4, FIT, *Pa*, and *Fn* performed best in discriminating CRC and AA from healthy controls, which was significantly better than any combination of two or three of the four markers. The pd-score obtained according to the four coefficients demonstrated an AUC of 0.958 for CRC and AA detection (Fig. 3F). Using the cutoff of 0.6, as determined by the Youden index method, the pd-score demonstrated a specificity of 91.37% and sensitivities of 95.38% for CRC and 70.83% for AA, respectively (Fig. 3G).

The pd-score performs significantly better than FIT in diagnosing AA and stage I–II CRC

FIT has been used for CRC diagnosis and screening for decades, so we assessed the superiority of the pd-score compared to FIT in discriminating CRC and AA from healthy subjects. The pd-score performed significantly better than FIT in discriminating CRC and AA from healthy subjects, with AUCs of 0.971 vs 0.892 for CRC and AUCs of 0.904 vs 0.721 for AA (Fig. 4A, B). The IDI analysis showed that the diagnostic capacity of the pd-score was 19.38% (95% CI 0.1545–0.2231, $P < 0.001$) higher than that of FIT in discriminating both CRC and AA from healthy subjects. Importantly, the pd-score was significantly more advantageous than FIT in diagnosing stage I–II CRC (sensitivity: 94.57% vs. 80.43%, $P = 0.002$) (Fig. 4C) and AA measuring greater than 1 cm (sensitivity: 71.11% vs. 48.89%, $P = 0.016$) (Fig. 4D), indicating a possible effective role of the pd-score in CRC screening.

Validation of the pd-score in the external validation cohort

Consistent with that in the SYSUCC cohort, the pd-scores were significantly higher in AA ($P < 0.001$) and CRC ($P < 0.001$) patients than in normal controls in the external validation cohort (Fig. 5A). Using the same cutoff value generated from the SYSUCC cohort, the pd-score and FIT identified 63 (94.03%) and 53 (79.10%) persons with CRC, 12 (80%) and 8 (53.33%) persons with AA, and 28 (80%) and 31 (88.57%) healthy subjects, respectively (Fig. 5B). The pd-score and FIT demonstrated AUCs of 0.952 and 0.838 for CRC and 0.874 and 0.710 for AA, respectively (Fig. 5C, D). The IDI analysis demonstrated a 24.55% (95% CI 0.1664–0.3254, $P < 0.001$) improvement in the pd-score compared to FIT for discriminating both CRC and AA from healthy subjects.

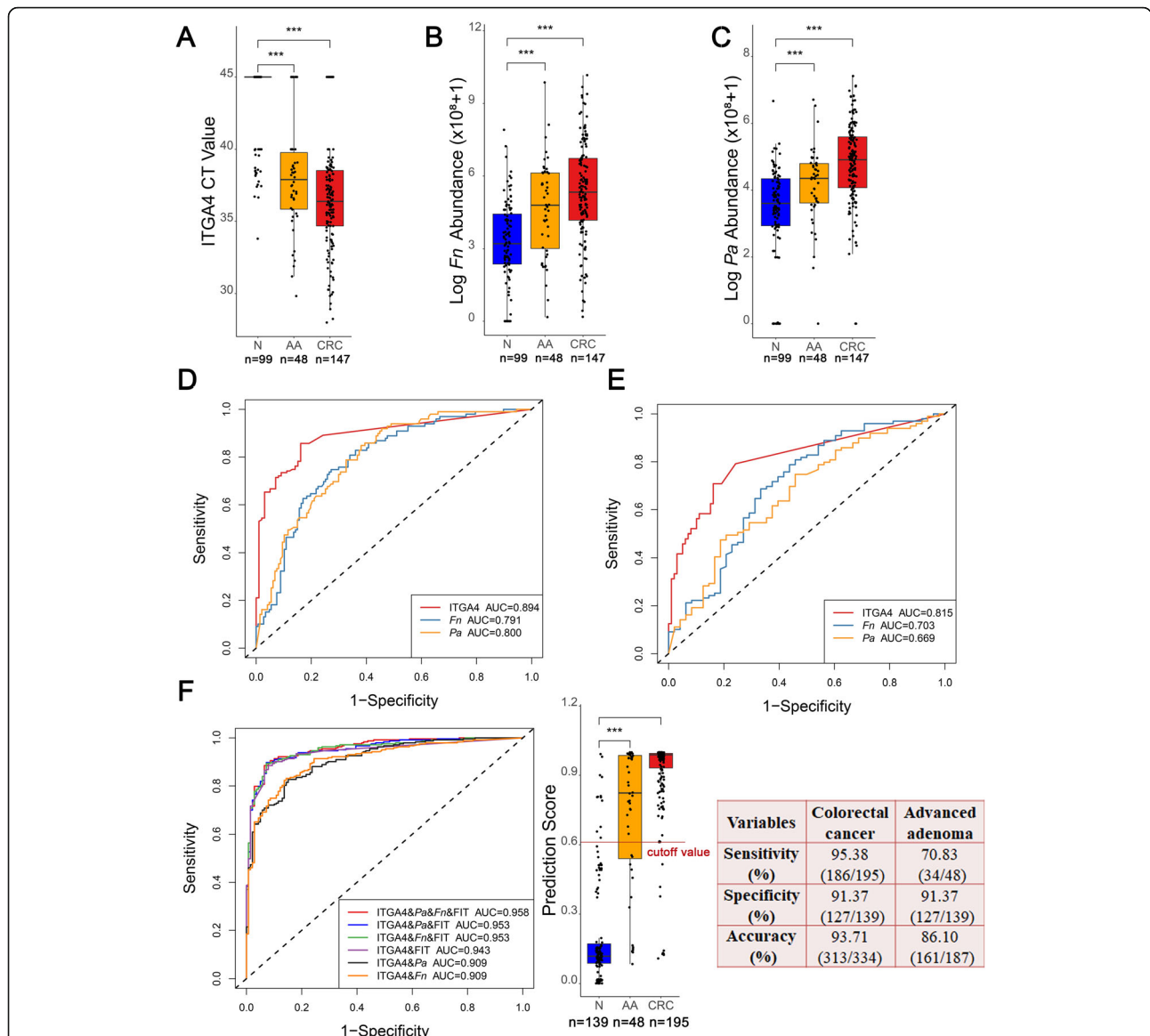


Fig. 3 Testing of the selected methylated DNA and bacterial markers and prediction model construction. **A–C** Methylation level of ITGA4 (**A**) and relative abundance of *Fn* (**B**) and *Pa* (**C**) in fecal samples collected from healthy subjects (N), AA patients and CRC patients. **D, E** ROC curves along with the corresponding AUCs of ITGA4, *Fn*, *Pa*, and FIT in the detection of CRC (**D**) and AA (**E**). **F** ROC curves along with the corresponding AUCs of different combinations of ITGA4, *Fn*, *Pa*, and FIT. **G** Prediction score in normal control and AA and CRC patients generated from the logistic regression model constructed by the combination of ITGA4, *Fn*, *Pa*, and FIT and the diagnostic performance of the pd-score using the cutoff value of 0.6 (determined by the Youden index method)

Validated screening effect of the pd-score in CRC screening cohort

Among the 749 participants who were eligible for evaluation in the CRC screening cohort, the pd-scores were significantly higher in CRC ($P < 0.001$), AA ($P < 0.001$), nonadvanced adenoma (non-AA) ($P = 0.012$) and polyps ($P = 0.035$) than in healthy subjects. The pd-scores were significantly decreased in postoperative CRC (CRC-post) patients compared with untreated patients ($P = 0.001$) (Fig. 6A). In addition, there was no significant difference

in the pd-scores between patients with IBD and healthy subjects ($P = 0.710$). The pd-scores of IBD and polyps were significantly lower than those of AA and CRC (all $P < 0.001$) (Additional file 1: Figure S2), indicating the clinical utility of the pd-score in differential diagnosis.

As the purpose of screening is to identify precancerous lesions and CRC, participants with non-AD and non-CRC were all considered controls in the following analysis. In this context, the pd-score and FIT demonstrated

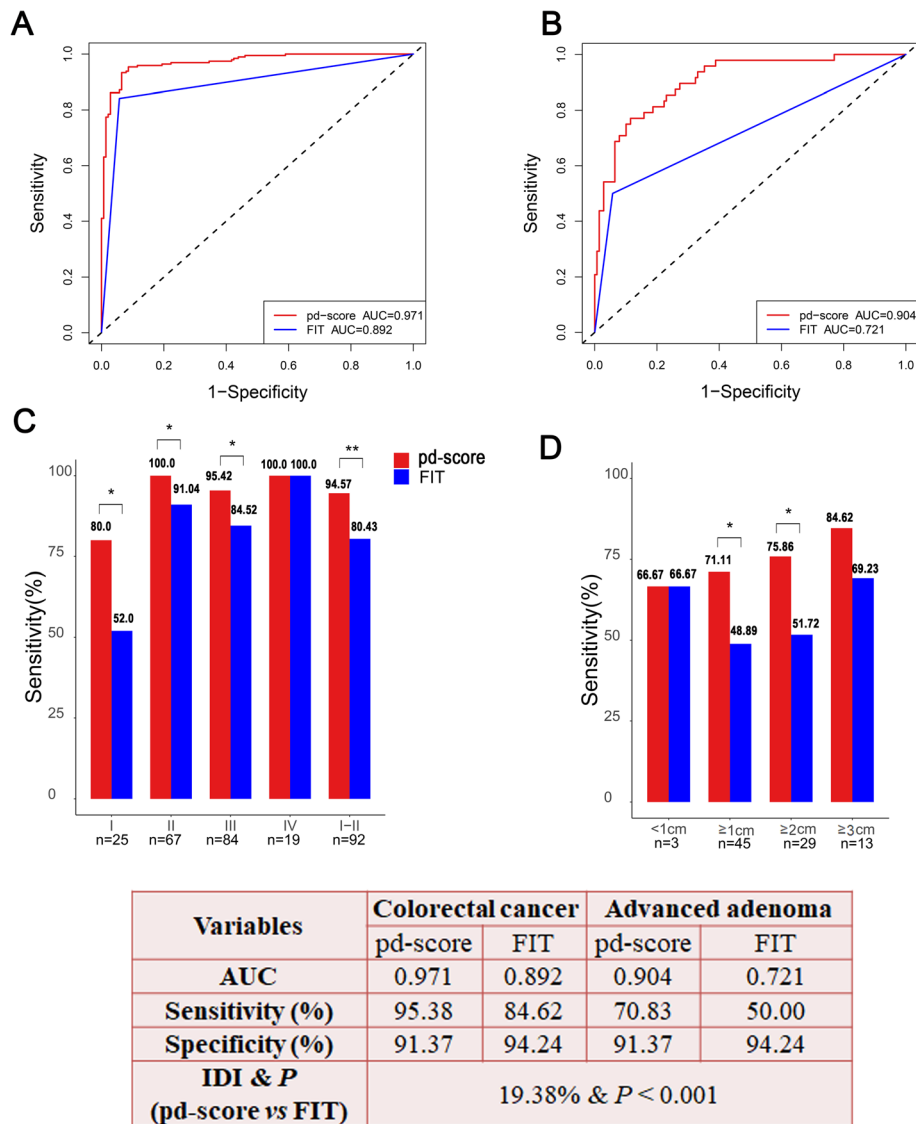


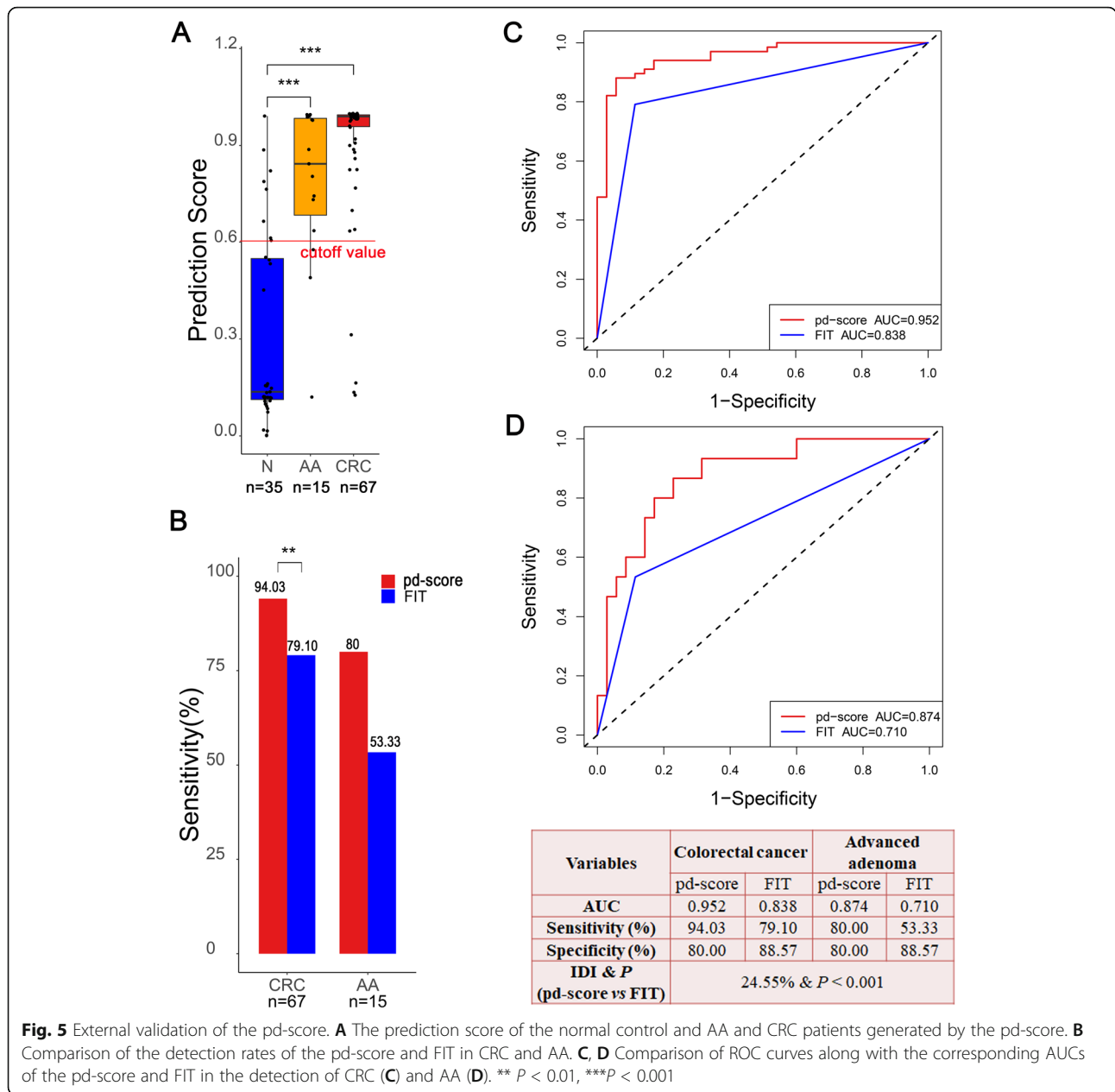
Fig. 4 Comparison of the pd-score with FIT in CRC and AA detection. **A, B** Comparison of ROC curves along with the corresponding AUCs of the pd-score and FIT in the detection of CRC (**A**) and AA (**B**). **C** Effect of cancer stage on the detection rates of CRC by the pd-score and FIT. **D** Effect of size on the detection rates of AA by the pd-score and FIT. NS, nonsignificant, * $P < 0.05$, ** $P < 0.01$

AUCs of 0.985 and 0.900 for CRC (Fig. 6C) and 0.795 and 0.587 for AA (Fig. 6D), respectively. Colonoscopy, the pd-score, and FIT identified 11, 11 (100%), and 11 (100%) persons with CRC; 104, 40 (38.46%), and 22 (21.15%) persons with AD; 80, 23 (28.75%), and 13 (16.25%) persons with non-AA; 24, 17 (70.83%), and 9 (37.5%) persons with AA; and 630, 485 (76.98%), and 503 (79.84%) normal controls, respectively. The pd-score showed a significantly higher sensitivity than FIT in diagnosing AA ($P = 0.013$, non-AA ($P = 0.004$), and AD ($P < 0.001$) (Fig. 6D). The IDI analysis showed that the pd-score improved the results by 41.19% (95% CI 0.3353–0.4884, $P < 0.001$) compared to FIT in discriminating CRC and AA from normal controls.

Discussion

The colorectal cancer incidence can be effectively decreased through the discovery and removal of advanced adenomas during screening, and CRC mortality can be remarkably reduced by its detection at an early, resectable stage. This study developed a prediction model using methylated DNA, bacterial markers, and FIT, which showed a high sensitivity for the detection of AA and early-stage CRC.

Strategies for CRC screening have shifted from one step (colonoscopy) to 2 steps, and a 2-step screening scenario is now the most acceptable strategy worldwide [41]. Colonoscopy is the gold standard for the diagnosis of CRC. However, it should be noted that, in addition to its invasive nature and the need for bowel preparation,

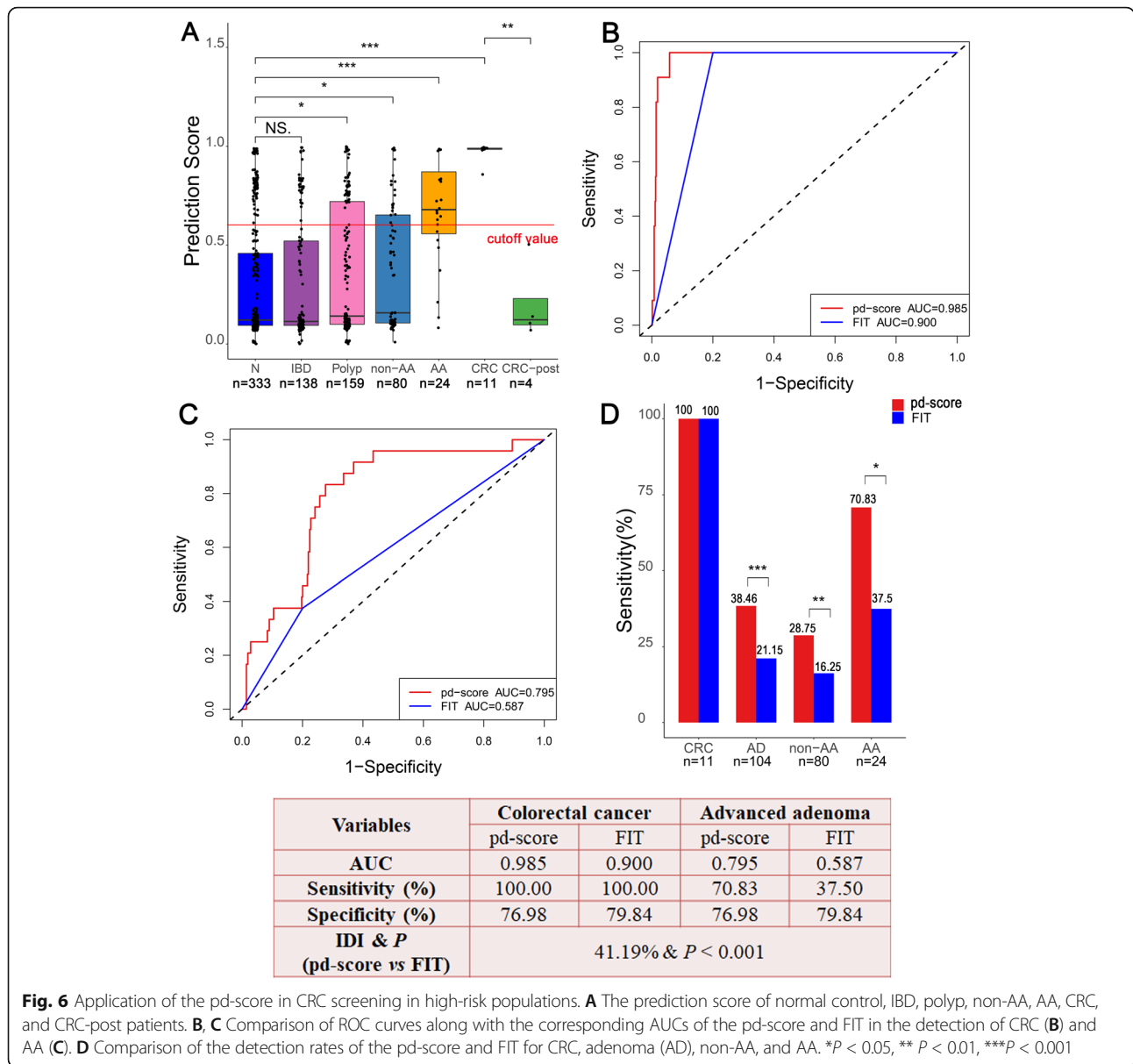


in countries with large populations, such as China, there are not enough doctors or colonoscopies to meet the needs of one-step screening.

FIT is now the most widely used CRC screening test but it is limited by a relatively low sensitivity for early-stage CRC and AA [42]. In addition, some factors may have impact on the FIT test and lead to false-positive results. A meta-analysis demonstrated that female was related to a significantly higher false-positive rate compared to males. People with medical history of peptic ulcer and anal fissure also tend to have significantly higher risk of false-positive FIT results possibly due to their high risk for bleeding. Using medications of NSAID

s (e.g., aspirin) may leads to higher false positive rate as well [43]. Another population-based screening program conducted in German revealed that older age and new diagnosis of IBD associated with higher risk of false-positive results [44]. Thus, it is rational to combine more specific approaches (sDNA) with FIT to improve the accuracy of CRC screening [14].

It is axiomatic that tests for cancer screening must be sensitive enough, since the primary aim of cancer screening is to filter out precancerous lesions and early-stage cancers [45]. Approximately 85% of CRC cases arise from AA; thus, effectively detecting AA patients is the key in CRC screening. Patients with early-stage CRC



have a nearly 90% 5-year survival rate, whereas it is only 14% for those diagnosed with distant disease [6]. Screening out CRC at early stages is crucial for effective surgical and therapeutic interventions. The pd-score showed a sensitivity of 94.57% for stage I–II CRC, which was close to the estimated sensitivity of colonoscopy [46–50]. In the CRC screening cohort, IBD patients were included for analysis and inclusion of these subjects could be closer to the actual scenarios that screening test were applied, since this cohort came from the community. The pd-scores of IBD were significantly different with CRC and were no difference with normal controls. In addition, we computed the specificity for the screening cohort with or without these IBD patients, respectively. The results showed that the specificity is stable

(without IBD: 76.83%, with IBD: 76.98%), indicating the combined predictor of pd-score can distinguish CRC from IBD patients.

The high sensitivity of the pd-score is due to the diagnostic complementarity of different types of markers. We performed a subgroup analysis and found high diagnostic complementarity of different markers in different stages and locations of tumors. We found that the detection rates of the four markers (FIT, ITGA4, *Fn*, and *Pa*) all tended to increase with stage (Additional file 1: Figure S3A). Compared with patients at advanced stages, AA and stage I patients are less likely to suffer from tumor hemorrhage; thus, the sensitivity of FIT is limited [14, 51]. In contrast to FIT, methylated ITGA4 is derived from exfoliated cells, and DNA methylation signatures

are considered to maintain during colorectal carcinogenesis from adenoma to cancer [52], which may help to explain the significantly higher sensitivity than FIT for AA and stage I patients. Several studies have reported that gut microbial dysbiosis is linked to the progression of colorectal neoplasia [53], and this study observed that the abundances of *Fn* and *Pa* changed with stage. *Fn* showed a sensitivity of 54.17% for AA patients, which contributed greatly to the pd-score of patients with AA. Regarding tumor location, we observed that FIT showed a higher sensitivity for distal CRC, while ITGA4, *Fn*, and *Pa* performed better in discriminating proximal CRC from healthy subjects (Additional file 1: FigureS3B). Hemoglobin from proximal tumors may degrade before reaching the anus, which probably leads to the inferiority of FIT for proximal CRC [54–56]. Differences in developmental origin and environmental mutagens lead to variations in molecular features and gut flora between proximal and distal CRC [57–59]. Previous studies have reported a positive CpG island methylator phenotype (CIMP), and the proportion of *Fn* gradually increased from the rectum to the proximal colon [60–62], which is consistent with our results. Proximal CRC, especially those at an early stage, is difficult to diagnose due to obscure clinical manifestations, and studies have found that the prognosis for patients with proximal CRC is worse than that for patients with distal CRC [63, 64]. Due to the complementarity of markers from different types, the pd-score shows high sensitivity for both AA and CRC, this is of great importance to improve the early diagnosis rate and the survival of proximal CRC patients.

Although the FIT-sDNA test has been accepted and recommended as a CRC screening method by the NCCN [15] and China guidelines [16], the grading of its recommendation is relatively low due to its lower level of supporting evidence. In the Chinese guidelines for screening CRC, the recommendation category of FIT is moderate, but it is low for the FIT-sDNA test [16]. An important reason is that there are few studies revealing the screening effect of the FIT-sDNA test in large populations. Here, in our study, in addition to being validated in multicenter hospital-based samples, our pd-score was also applied in a CRC screening cohort. In multicenter samples, the diagnostic performance of the pd-score improved 24.55% compared with FIT in discriminating CRC and AA from controls. When applied in a screening setting, it achieved a 41.19% improvement in the pd-score compared to FIT in distinguishing CRC and AA from controls. Our study showed that combining FIT and altered fecal DNA yielded higher single-application diagnostic performance than a commercial FIT for both CRC and AA, providing evidence for the application of the FIT-sDNA test in CRC screening and early diagnosis.

This study had obvious strengths. First, this is the first FIT-sDNA test validated in a CRC screening cohort in

China. Second, the pd-score constructed in this study showed high sensitivity for advanced adenoma and was well validated in both the multicenter and screening cohorts. Third, the evaluation of the pd-score was performed in comparison to a commercial FIT.

Nevertheless, there are a few limitations to be considered for this study. First, the number of AA samples in the model construction is relatively small, which may lead to a limited power to estimate the sensitivity. However, the sensitivity of the pd-score for AA in the external validation cohort and the CRC screening cohort was more or less in agreement with that in the model construction cohort, indicating the reliability and accuracy of the pd-score in discriminating AA patients from healthy controls. Second, whether the pd-score will maintain good diagnostic performance in other ethnic populations remains unknown. In the end, the biomarkers of our study were selected based on a candidate strategy, thus some effective ones could be missed. Although the validate samples have confirmed its screening efficiency and high potential in practical application, future biomarkers with better performance are still needed to improve the screening efficiency

Conclusion

The current study provided a noninvasive and convenient approach with high diagnostic accuracy for advanced adenomas and early-stage CRC. Importantly, we constructed a pd-score that was generated from a large case-control cohort and was well validated in an external validation cohort and CRC screening cohort. The capacity for accurate identification of both precancerous lesions and curable-stage CRC will allow the pd-score to be applied widely as a diagnostic method and a screening test.

Abbreviations

AA: Advanced adenoma; AD: Adenoma; AJCC: American Joint Committee on Cancer; AUC: Area under the curve; *Ch*: *Clostridium hathewayi*; CIMP: CpG island methylator phenotype; CRC: Colorectal cancer; FDA: Food and Drug Administration; FIT: Fecal immunochemical test; *Fn*: *Fusobacterium nucleatum*; FPB: Fecal preservative buffer; IBD: Inflammatory bowel disease; IDI: Integrated discrimination improvement; *Pa*: *Peptostreptococcus anaerobius*; *Pm*: *Parvimonas micra*; post-CRC: Postoperative CRC; sDNA: Stool DNA; *Sm*: *Solobacterium moorei*

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-021-02123-0>.

Additional file 1: Figure S1. Effect of the FPB on human genome DNA integrity and bacterial diversity. **Figure S2.** The prediction score of IBD, polyp, non-AA, AA and CRC patients. **Figure S3.** Detection rates of ITGA4, *Fn*, *Pa*, FIT and pd-score on different stage and position.

Additional file 2: Table S1. Clinical characteristics of the SYSUCC cohort. **Table S2.** Clinical characteristics of the external validation cohort. **Table S3.** Clinical characteristics of the SZ CRC screening cohort. **Table S4.** Primer and probe sequences for quantitative real-time PCR of methylated DNA markers. **Table S5.** Primer sequences for quantitative real-time PCR of bacterial markers.

Acknowledgements

We thank Wan-li Liu (Sun Yat-sen University Cancer Center) and Zun-fu Ke (the First affiliated hospital of Sun Yat-sen University) for their help with the sample collection. We thank Kai-wen Zhang for her assistance with the manuscript editing.

Authors' contributions

WJ conceived, designed, and directed the study. LC, JZ, TZ, and ZW developed the methodology. LC, XP, WX, RZ, GL, XT, FW, KC, SZ, LZ, and PD collected the samples. LC, TW, DW, and YH performed the analyses and interpretation of data. YL provided technical or material support. LC wrote the manuscript. WJ, YH, and WX critically reviewed the manuscript. All authors read and approved the final manuscript. The order of the co-first authors was assigned based on the relative contributions of the individuals.

Funding

This work was supported by research funds from the National Key Research and Development Program of China (2016YFC1302700), Health and Family Commission of Shenzhen Municipality Funded Project (SZGW2017011), and Nanshan Science and Technology Innovation Bureau (2017007, 2019054 and 2020097).

Availability of data and materials

The authenticity of this article has been validated by uploading the key raw data onto the Research Data Deposit public platform (www.researchdata.org.cn), with the approval RDD number as RDDA2021001989.

Declarations

Ethics approval and consent to participate

This project was approved by the Research Ethics Committee of the SYSUCC, and prior consent was obtained from all participants for experimentation with human samples. The CRC screening project was approved by the Ethics Committee of the Shenzhen Nanshan Center for Chronic Disease Control, and prior consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Oncology in South China Guangzhou, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, People's Republic of China. ²Department of Radiation Oncology, Affiliated Hospital of Qingdao University, Qingdao, People's Republic of China. ³Shenzhen Nanshan Center for Chronic Disease Control, Shenzhen, People's Republic of China. ⁴Department of Endoscopy and Laser, Sun Yat-Sen University Cancer Center, Guangzhou, People's Republic of China. ⁵Biobank of Sun Yat-sen University Cancer Center, Guangzhou, People's Republic of China. ⁶School of Public Health, Sun Yat-Sen University, Guangzhou, People's Republic of China. ⁷Department of Radiation Oncology, Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, Guangdong, People's Republic of China. ⁸Department of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin Key Laboratory of Breast Cancer Prevention and Therapy, Ministry of Education, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, People's Republic of China. ⁹Department of Laboratory Medicine, First Affiliated Hospital, Sun Yat-sen University, Guangzhou, People's Republic of China. ¹⁰Department of Colorectal Surgery, Sun Yat-sen University Cancer Center, Guangzhou, People's Republic of China.

Received: 3 June 2021 Accepted: 13 September 2021

Published online: 25 October 2021

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020; 70(1):7–30. <https://doi.org/10.3322/caac.21590>.

- Ferlizza E, Solmi R, Sgarzi M, Ricciardiello L, Lauriola M. The Roadmap of Colorectal Cancer Screening. *Cancers*. 2021;13(5):1101. <https://doi.org/10.3390/cancers13051101>.
- Zhang S, Sun K, Zheng R, Zeng H, Wang S, Chen R, et al. Cancer incidence and mortality in China 2015. *J Natl Cancer Center*. 2020;1(1):–11. <https://doi.org/10.1016/j.jncc.2020.12.001>.
- Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet*. 2019;394(10207):1467–80.
- Meester RG, Doubeni CA, Lansdorp-Vogelaar I, Goede SL, Levin TR, Quinn VP, et al. Colorectal cancer deaths attributable to nonuse of screening in the United States. *Ann Epidemiol*. 2015;25(3):208–13 e1.
- Provenzale D, Ness RM, Llor X, Weiss JM, Abbadesse B, Cooper G, et al. NCCN Guidelines Insights: Colorectal Cancer Screening, Version 2.2020. *J Natl Compr Cancer Netw*. 2020;18(10):1312–20. <https://doi.org/10.6004/jnccn.2020.0048>.
- Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. *Gut*. 2015;64(5):784–90. <https://doi.org/10.1136/gutjnl-2014-307508>.
- 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. <https://doi.org/10.7326/M13-1484>.
- Robertson DJ, Lee JK, Boland CR, Dominitz JA, Giardiello FM, Johnson DA, 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):1217–37 e3. <https://doi.org/10.1053/j.gastro.2016.08.053>.
- Chao S, Ying J, Liew G, Marshall W, Liew CC, Burakoff R. Blood RNA biomarker panel detects both left- and right-sided colorectal neoplasms: a case-control study. *J Exp Clin Cancer Res*. 2013;32(1):44. <https://doi.org/10.1186/1756-9966-32-44>.
- Ciarloni L, Ehrensberger SH, Imaizumi N, Monnier-Benoit S, Nichita C, Myung SJ, et al. Development and clinical validation of a blood test based on 29-gene expression for early detection of colorectal cancer. *Clin Cancer Res*. 2016;22(18):4604–11. <https://doi.org/10.1158/1078-0432.CCR-15-2057>.
- Ferlizza E, Solmi R, Miglio R, Nardi E, Mattei G, Sgarzi M, et al. Colorectal cancer screening: Assessment of CEACAM6, LGALS4, TSPAN8 and COL1A2 as blood markers in faecal immunochemical test negative subjects. *J Adv Res*. 2020;24:99–107. <https://doi.org/10.1016/j.jare.2020.03.001>.
- Peterse EFP, Meester RGS, de Jonge L, Omidvari AH, Alarid-Escudero F, Knudsen AB, et al. Comparing the cost-effectiveness of innovative colorectal cancer screening tests. *J Natl Cancer Inst*. 2021;113(2):154–61. <https://doi.org/10.1093/jnci/djaa103>.
- Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med*. 2014;370(14):1287–97. <https://doi.org/10.1056/NEJMoa1311194>.
- Provenzale D, Gupta S, Ahnen DJ, Markowitz AJ, Chung DC, Mayer RJ, et al. NCCN guidelines insights: colorectal cancer screening, Version 1.2018. *J Natl Compr Cancer Netw*. 2018;16(8):939–49. <https://doi.org/10.6004/jnccn.2018.0067>.
- National Cancer Center CEGoTDoCGfSED Early Treatment of Colorectal C. China guideline for the screening, early detection and early treatment of colorectal cancer (2020, Beijing). *Zhonghua Zhong Liu Za Zhi*. 2021;43(1): 16–38.
- Nguyen LH, Goel A, Chung DC. Pathways of colorectal carcinogenesis. *Gastroenterology*. 2020;158(2):291–302. <https://doi.org/10.1053/j.gastro.2019.08.059>.
- Jung G, Hernández-Illán E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol*. 2020;17(2):111–30. <https://doi.org/10.1038/s41575-019-0230-y>.
- Yan W, Guo M. Epigenetics of colorectal cancer. *Methods Mol Biol (Clifton, NJ)*. 2015;1238:405–24.
- Kanai Y, Hirohashi S. Alterations of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state. *Carcinogenesis*. 2007; 28(12):2434–42. <https://doi.org/10.1093/carcin/bgm206>.
- Mu J, Huang Y, Cai S, Li Q, Song Y, Yuan Y, et al. Plausibility of an extensive use of stool DNA test for screening advanced colorectal neoplasia. *Clin Chim Acta*. 2020;501:42–7.
- Rengucci C, De Maio G, Menghi M, Scarpi E, Guglielmo S, Fusaroli P, et al. Improved stool DNA integrity method for early colorectal cancer diagnosis.

- Cancer Epidemiol Biomarkers Prev. 2014;23(11):2553–60. <https://doi.org/10.1158/1055-9965.EPI-14-0379>.
23. Calistri D, Rengucci C, Casadei Gardini A, Frassinetti GL, Scarpi E, Zoli W, et al. Fecal DNA for noninvasive diagnosis of colorectal cancer in immunochemical fecal occult blood test-positive individuals. *Cancer Epidemiol Biomarkers Prev.* 2010;19(10):2647–54. <https://doi.org/10.1158/1055-9965.EPI-10-0291>.
 24. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res.* 2012;22(2):299–306. <https://doi.org/10.1101/gr.126516.1111>.
 25. Feng Q, Liang S, Jia H, Stadlmayr A, Tang L, Lan Z, et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat Commun.* 2015;6(1):6528. <https://doi.org/10.1038/ncomms7528>.
 26. Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut.* 2017;66(1):70–8. <https://doi.org/10.1136/gutjnl-2015-309800>.
 27. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med.* 2019;25(4):667–78. <https://doi.org/10.1038/s41591-019-0405-7>.
 28. Guo S, Li L, Xu B, Li M, Zeng Q, Xiao H, et al. A simple and novel fecal biomarker for colorectal cancer: ratio of Fusobacterium nucleatum to probiotics populations, based on their antagonistic effect. *Clin Chem.* 2018; 64(9):1327–37. <https://doi.org/10.1373/clinchem.2018.289728>.
 29. Jalanka J, Salonen A, Salojärvi J, Ritari J, Immonen O, Marciari L, et al. Effects of bowel cleansing on the intestinal microbiota. *Gut.* 2015;64(10):1562–8. <https://doi.org/10.1136/gutjnl-2014-307240>.
 30. Chen H, Li N, Ren J, Feng X, Lyu Z, Wei L, et al. Participation and yield of a population-based colorectal cancer screening programme in China. *Gut.* 2019;68(8):1450–7. <https://doi.org/10.1136/gutjnl-2018-317124>.
 31. Chen WD, Han ZJ, Skoletsky J, Olson J, Sah J, Myeroff L, et al. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J Natl Cancer Inst.* 2005;97(15):1124–32. <https://doi.org/10.1093/jnci/dji204>.
 32. Itzkowitz S, Brand R, Jandorf L, Durke K, Millholland J, Rabeneck L, et al. A simplified, noninvasive stool DNA test for colorectal cancer detection. *Am J Gastroenterol.* 2008;103(11):2862–70. <https://doi.org/10.1111/j.1572-0241.2008.02088.x>.
 33. Niu F, Wen J, Fu X, Li C, Zhao R, Wu S, et al. Stool DNA test of methylated syndecan-2 for the early detection of colorectal neoplasia. *Cancer Epidemiol Biomark Prev.* 2017;26(9):1411–9. <https://doi.org/10.1158/1055-9965.EPI-17-0153>.
 34. Ausch C, Kim YH, Tsuchiya KD, Dzieciatkowski S, Washington MK, Paraskeva C, et al. Comparative analysis of PCR-based biomarker assay methods for colorectal polyp detection from fecal DNA. *Clin Chem.* 2009;55(8):1559–63. <https://doi.org/10.1373/clinchem.2008.122937>.
 35. Lind GE, Danielsen SA, Ahlquist T, Merok MA, Andresen K, Skotheim RI, et al. Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. *Mol Cancer.* 2011;10(1):85. <https://doi.org/10.1186/1476-4598-10-85>.
 36. Liang JQ, Li T, Nakatsu G, Chen YX, Yau TO, Chu E, et al. A novel faecal Lachnospirillum marker for the non-invasive diagnosis of colorectal adenoma and cancer. *Gut.* 2020;69(7):1248–57. <https://doi.org/10.1136/gutjnl-2019-318532>.
 37. Wong SH, Kwong TNY, Chow TC, Luk AKC, Dai RZW, Nakatsu G, et al. Quantitation of faecal Fusobacterium improves faecal immunochemical test in detecting advanced colorectal neoplasia. *Gut.* 2017;66(8):1441–8. <https://doi.org/10.1136/gutjnl-2016-312766>.
 38. Nani BD, Lima PO, Marcondes FK, Groppo FC, Rolim GS, Moraes AB, et al. Changes in salivary microbiota increase volatile sulfur compounds production in healthy male subjects with academic-related chronic stress. *PLoS One.* 2017; 12(3):e0173686. <https://doi.org/10.1371/journal.pone.0173686>.
 39. Rath S, Heidrich B, Pieper DH, Vital M. Uncovering the trimethylamine-producing bacteria of the human gut microbiota. *Microbiome.* 2017;5(1):54. <https://doi.org/10.1186/s40168-017-0271-9>.
 40. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837–45. <https://doi.org/10.2307/2531595>.
 41. Young GP, Rabeneck L, Winawer SJ. The Global Paradigm Shift in Screening for Colorectal Cancer. *Gastroenterology.* 2019;156(4):843–51 e2. <https://doi.org/10.1053/j.gastro.2019.02.006>.
 42. Quintero E, Castells A, Bujanda L, Cubiella J, Salas D, Lanás A, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. *N Engl J Med.* 2012;366(8):697–706. <https://doi.org/10.1056/NEJMoa1108895>.
 43. de Klerk CM, Vendrig LM, Bossuyt PM, Dekker E. Participant-related risk factors for false-positive and false-negative fecal immunochemical tests in colorectal cancer screening: systematic review and meta-analysis. *Am J Gastroenterol.* 2018;113(12):1778–87. <https://doi.org/10.1038/s41395-018-0212-7>.
 44. Amitay EL, Cuk K, Niedermaier T, Weigl K, Brenner H. Factors associated with false-positive fecal immunochemical tests in a large German colorectal cancer screening study. *Int J Cancer.* 2019;144(10):2419–27. <https://doi.org/10.1002/ijc.31972>.
 45. Ahlquist DA. Multi-target stool DNA test: a new high bar for noninvasive screening. *Dig Dis Sci.* 2015;60(3):623–33. <https://doi.org/10.1007/s10620-014-3451-5>.
 46. Bressler B, Paszat LF, Vinden C, Li C, He J, Rabeneck L. Colonoscopic miss rates for right-sided colon cancer: a population-based analysis. *Gastroenterology.* 2004;127(2):452–6. <https://doi.org/10.1053/j.gastro.2004.05.032>.
 47. Patel SG, Ahnen DJ. Prevention of interval colorectal cancers: what every clinician needs to know. *Clin Gastroenterol Hepatol.* 2014;12(1):7–15. <https://doi.org/10.1016/j.cgh.2013.04.027>.
 48. Cooper GS, Xu F, Barnholtz Sloan JS, Schluchter MD, Koroukian SM. Prevalence and predictors of interval colorectal cancers in medicare beneficiaries. *Cancer.* 2012;118(12):3044–52. <https://doi.org/10.1002/ncr.26602>.
 49. Singh H, Nugent Z, Mahmud SM, Demers AA, Bernstein CN. Predictors of colorectal cancer after negative colonoscopy: a population-based study. *Am J Gastroenterol.* 2010;105(3):663–73; quiz 74. <https://doi.org/10.1038/ajg.2009.650>.
 50. Samadpour NJ, Curtin K, Tuohy TM, Pappas L, Boucher K, Provenzale D, et al. Characteristics of missed or interval colorectal cancer and patient survival: a population-based study. *Gastroenterology.* 2014;146(4):950–60. <https://doi.org/10.1053/j.gastro.2014.01.013>.
 51. Digby J, Fraser CG, Carey FA, McDonald PJ, Strachan JA, Diamant RH, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. *J Clin Pathol.* 2013;66(5):415–9. <https://doi.org/10.1136/jclinpath-2013-201445>.
 52. Bormann F, Rodríguez-Paredes M, Lasitschka F, Edelmann D, Musch T, Benner A, et al. Cell-of-origin DNA methylation signatures are maintained during colorectal carcinogenesis. *Cell Rep.* 2018;23(11):3407–18. <https://doi.org/10.1016/j.celrep.2018.05.045>.
 53. Nakatsu G, Li X, Zhou H, Sheng J, Wong SH, Wu WK, et al. Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nat Commun.* 2015; 6(1):8727. <https://doi.org/10.1038/ncomms9727>.
 54. Zorzi M, Hassan C, Capodaglio G, Narne E, Turrin A, Baracco M, et al. Divergent long-term detection rates of proximal and distal advanced neoplasia in fecal immunochemical test screening programs: a retrospective cohort study. *Ann Intern Med.* 2018;169(9):602–9. <https://doi.org/10.7326/M18-0855>.
 55. Chiu HM, Jen GH, Wang YW, Fann JC, Hsu CY, Jeng YC, et al. Long-term effectiveness of faecal immunochemical test screening for proximal and distal colorectal cancers. *Gut.* 2021. [gutjnl-2020-322545](https://doi.org/10.1136/gutjnl-2020-322545); <https://doi.org/10.1136/gutjnl-2020-322545>.
 56. Wong MC, Ching JY, Chan VC, Lam TY, Shum JP, Luk AK, et al. Diagnostic accuracy of a qualitative fecal immunochemical test varies with location of neoplasia but not number of specimens. *Clin Gastroenterol Hepatol.* 2015; 13(8):1472–9. <https://doi.org/10.1016/j.cgh.2015.02.021>.
 57. Wang L, Lo CH, He X, Hang D, Wang M, Wu K, et al. Risk factor profiles differ for cancers of different regions of the colorectum. *Gastroenterology.* 2020; 159(1):241–56 e13. <https://doi.org/10.1053/j.gastro.2020.03.054>.
 58. Flemer B, Lynch DB, Brown JM, Jeffery IB, Ryan FJ, Claesson MJ, et al. Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut.* 2017;66(4):633–43. <https://doi.org/10.1136/gutjnl-2015-309595>.
 59. Missiaglia E, Jacobs B, D'Ario G, Di Narzo AF, Sonesson C, Budinska E, et al. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. *Ann Oncol.* 2014;25(10):1995–2001. <https://doi.org/10.1093/annonc/mdu275>.

60. Yamauchi M, Morikawa T, Kuchiba A, Imamura Y, Qian ZR, Nishihara R, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut*. 2012;61(6):847–54. <https://doi.org/10.1136/gutjnl-2011-300865>.
61. Arain MA, Sawhney M, Sheikh S, Anway R, Thyagarajan B, Bond JH, et al. CIMP status of interval colon cancers: another piece to the puzzle. *Am J Gastroenterol*. 2010;105(5):1189–95. <https://doi.org/10.1038/ajg.2009.699>.
62. Mima K, Cao Y, Chan AT, Qian ZR, Nowak JA, Masugi Y, et al. Fusobacterium nucleatum in colorectal carcinoma tissue according to tumor location. *Clin Transl Gastroenterol*. 2016;7(11):e200. <https://doi.org/10.1038/ctg.2016.53>.
63. Rahbari NN, Carr PR, Jansen L, Chang-Claude J, Weitz J, Hoffmeister M, et al. Time of metastasis and outcome in colorectal cancer. *Ann Surg*. 2019;269(3):494–502. <https://doi.org/10.1097/SLA.0000000000002564>.
64. Le H, Ziogas A, Taylor TH, Lipkin SM, Zell JA. Survival of distinct Asian groups among colorectal cancer cases in California. *Cancer*. 2009;115(2):259–70. <https://doi.org/10.1002/cncr.24034>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

