# Diagnostic accuracy of fecal calprotectin for screening patients with colorectal cancer: A meta-analysis

*Xiaohua Ye[1], Jiaping Huai[2], Jin Ding[1]*

[1]Department of Gastroenterology, Jinhua Hospital of Zhejiang University, Jinhua, Zhejiang, China
[2]Department of Critical Care Medicine, Jinhua Hospital of Zhejiang University, Jinhua, Zhejiang, China

***Cite this article as:*** *Ye X, Huai J, Ding J. Diagnostic accuracy of fecal calprotectin for screening patients with colorectal cancer: A meta-analysis. Turk J Gastroenterol DOI: 10.5152/tjg.2018.17606*

## ABSTRACT

***Background/Aims:*** *Fecal calprotectin (FC) is reported to have a broad diagnostic accuracy for colorectal cancer (CRC). Therefore, we explored the diagnostic value of FC for CRC using meta-analytical techniques to substantiate the assertion.*

***Materials and Methods:*** *An electronic search of the MEDLINE and Embase databases was conducted to identify studies that assessed the diagnostic accuracy of FC for CRC. The sensitivities and specificities of the eligible studies were summarized using a bivariable random-effects model.*

***Results:*** *In total, 20 studies were included in the final analysis. The pooled sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of FC for CRC detection were 0.83 (95% confidence interval (CI), 0.77-0.88), 0.61 (95% CI, 0.54-0.68), 2.15 (95% CI, 1.82-2.55), and 0.28 (95% CI, 0.21-0.37), respectively. The overall diagnostic odds ratio of FC for CRC was 7.76 (95% CI, 5.41-11.12) with an area under the curve of 0.81 (95% CI, 0.77-0.84), whereas the diagnostic value of FC for colorectal adenoma was relatively inferior (area under the curve, 0.55; 95% CI, 0.51-0.59; diagnostic odds ratio, 1.27; 95% CI, 0.91-1.78).*

***Conclusion:*** *The results imply that the FC test, as currently implemented, cannot be recommended for CRC detection.*

***Keywords:*** *Meta-analysis, calprotectin, colorectal cancer, diagnostic accuracy*

## INTRODUCTION

Colorectal cancer (CRC) is the third most common malignancy worldwide and accounts for >10% of all cancer deaths (1). Based on the guidelines of the National Institute for Health and Care Excellence, serious pathology is predicted by clinical manifestations, including rectal bleeding, weight loss, and a change in bowel habits (1). However, these symptoms are not specific to CRC and hence have a rather low positive predictive value of 3%-4%, thereby precluding their utility for establishing a diagnosis of CRC (1). Although colonoscopy is considered the gold standard for early detection of CRC and its precursors (adenoma), this invasive and expensive method involves patient discomfort and potential complications (2). Therefore, a non-invasive screening method is needed to identify patients suspected of having CRC who otherwise would require colonoscopy.

Fecal calprotectin (FC) is a mucosal neutrophil degradation product, the presence of which appropriately correlates with bowel inflammation (2,3). FC is also a well-established biomarker of inflammatory bowel disease (1). Moreover, FC is associated other bowel diseases, including CRC, adenoma, diverticulitis, and infectious diarrhea (4).

Colorectal cancer is associated with a local, acute inflammatory reaction of varying intensity. It is hypothesized that the local release of the chemotactic factor leads to the recruitment of neutrophils to the tumor site (4). A variety of inflammatory cells primarily along the invasive margin infiltrate human CRC tissue. Calprotectin activity is frequently detected in granulocytes and macrophages in CRC, and an increased numbers of granulocytes have been found in the feces of patients with CRC, likely because of shedding from the ulcerated tumor (5). Furthermore, it has been hypothesized that circulating leukocytes may actively migrate through neoplastic tissues in response to intraluminal antigens (5).

Numerous studies have evaluated the diagnostic accuracy of FC for the early detection of CRC. Despite the large body of evidence, conflicting results have raised concerns about the predictive value of FC for CRC detection across

***ORCID IDs of the authors:*** *X.Y. 0000-0002-8358-4093; J.H. 0000-0001-6778-6181; J.D. 0000-0002-5543-7969.*

*Corresponding Author:* ***Jin Ding; dingjin@163.com***
*Received:* ***September 30, 2017*** *Accepted:* ***January 29, 2018*** *Available online date:* ***May 30, 2018***

*DOI:* ***10.5152/tjg.2018.17606***

various patient populations. Indeed, a meta-analysis published in 2007 has revealed that FC is not suitable as a screening test for CRC because of its relatively low diagnostic accuracy (6). After a decade, many new studies have been published regarding the diagnostic accuracy of FC for CRC. Thus, we conducted a systematic review and meta-analysis to update the current evidence.

## MATERIALS AND METHODS

### Search Strategy

For our meta-analysis, we followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses and guidelines from the Cochrane Diagnostic Test Accuracy Working Group (7). A computerized search of MEDLINE (from January 1, 1966 to December 31, 2016) and EMBASE (from January 1, 1974 to January 31, 2017) was performed to identify potentially relevant articles. The search was conducted using the following keywords: calprotectin ("calprotectin", "leukocyte L1 antigen complex", and "calgranulin"), colorectal ("colon", "rectal", "intestinal", and "bowel"), and cancer ("malignancy", "neoplasm", "tumor", and "adenoma"). In addition, the bibliographies from these potential articles were manually searched to identify additional studies. Citations were limited to those published in English. This work was approved by the Ethics Committee of our hospital. Informed consent was waived off because it was a meta-analysis.

### Study Selection

Two reviewers (X.Y. and J.H.) independently reviewed potentially relevant articles for eligibility and inclusion. Studies were included if they met the following inclusion criteria: (1) prospective design published in a manuscript form; (2) patients for whom FC was used to detect CRC or colorectal adenoma; (3) an appropriate reference standard (endoscopy, radiology, or pathology) was included; (4) absolute numbers of true-positive, false-negative, true-negative, and false-positive observations for CRC and colorectal adenoma were reported or if data were sufficient to construct a 2×2 contingency table. Case reports, editorials, review articles, and clinical guidelines were excluded. Disagreements were resolved by a consensus.

### Data Extraction and Quality Assessment

A custom-made standardized form was used for data extraction. The following data were extracted for each eligible study: the surname of the first author, publication year, location of the study population, study design, sample size (the number of patients with CRC or adenoma and the total number of subjects screened), details of the FC test (FC assay method, cut-off value), and a reference standard. The results were transformed by multiplying each value by a factor of 5, as previously confirmed, because some studies used a conventional assay to measure FC in terms of milligrams per liter (6-8).

The methodological qualities of the studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (9). QUADAS-2 consists of four key categories: patient selection, index test, reference standard, and flow and timing. Each category was assessed in terms of the risk of bias, and the first three categories were also considered to determine applicability.

### Data Synthesis and Statistical Analysis

The 2×2 tables (numbers of true-positives, false-positives, true-negatives, and false-negatives) were constructed in accordance with the data from the included studies. The pooled estimates of sensitivity and specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) were obtained using a bivariate random-effects model (10). Also, a hierarchical summary receiver-operating characteristic (HSROC) curve was produced to present the summary estimates of sensitivities and specificities along with their corresponding 95% confidence interval (CI) and prediction region. Moreover, the area under the HSROC curve (AUC) was calculated (11). The AUC and DOR values were used to evaluate the diagnostic efficacy of FC for the CRC and colorectal adenoma groups. A DOR of 1 implies that the test is incapable of determining whether a patient has a colorectal neoplasm. A higher value indicates better test performance.

Heterogeneity was evaluated using the Q-statistic and quantified using $I^2$. For the Q test, a p value of <0.10 was considered to imply heterogeneity. The value of $I^2$ represents the proportion of total variation contributed by the between-study variation. The source of heterogeneity was explored using a threshold analysis, subgroup analysis, and meta-regression. Fagan plots were applied to illustrate the relationship among the prior test probability, PLR/NLR, and posterior test probability, and the Deek's test was used to evaluate publication bias (12). All statistical analyses were conducted using the STATA software version 12.0 (College Station; Texas, USA). A p value of <0.05 was considered to reflect statistical significance.

## RESULTS

### Study Selection and Characteristics

Our literature search identified 213 articles, of which 181 were then excluded upon an examination of the titles or abstracts because they were reviews, experimental studies, meta-analyses, comments, or other irrelevant articles (diagnostic studies of other biomarkers, letters, editorials, and consensus statements). Of the remaining 32 articles, 12 were subsequently excluded from the meta-analysis because two provided insufficient data to construct a 2×2 table, five did not focus on the diagnostic study of interest, one did not have a control group, three were not prospective design, and one reported other cancers not relevant to CRC. As a result, 20 studies were identified for inclusion in our meta-analysis (Figure 1) (1-4,13-28).

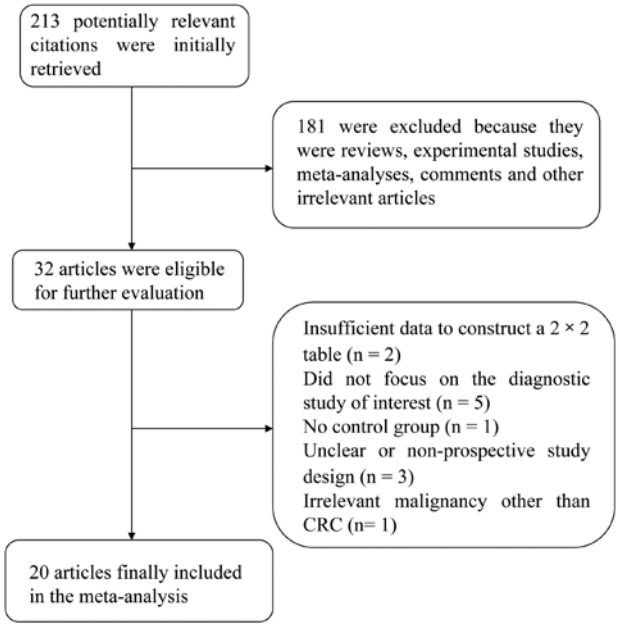


**Figure 1.** Flow chart for the selection of published studies for inclusion in our analysis

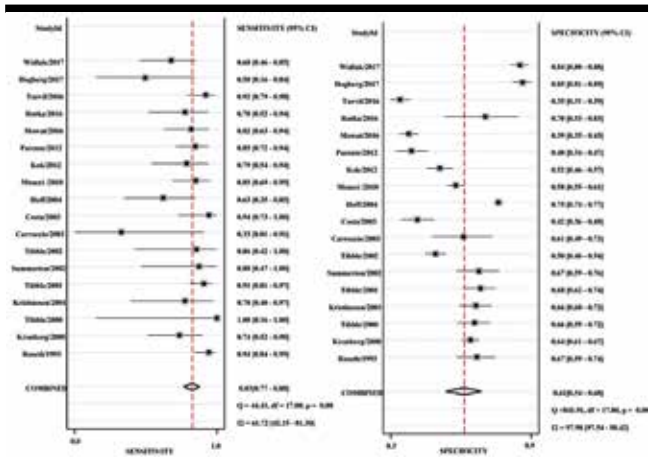

**Figure 2.** Forrest plots of sensitivity and specificity data for FC for CRC detection

Baseline characteristics of the 20 eligible studies are summarized in Table 1. These studies were published between 1993 and 2017 and were performed in seven regions. Of the 20 studies, 18 were conducted in Europe and two in the USA. The number of participants in each study ranged from 80 to 2,321 and included a total of 427 cases with a diagnosis of CRC and 1,806 cases with a diagnosis of adenoma (one study presented 40 cases that described polyps) (13). Overall, six studies used the conventional FC assay, 12 used the new assay , one used both assay, and one did not specify the detection method (1-4,13-28). Based on the results from QUADAS-2, the study bias and applicability outcomes were assessed; the results are shown in supplementary Figure S1. Of the 20 studies, 11 were judged as high-risk in one or more of the four key categories.

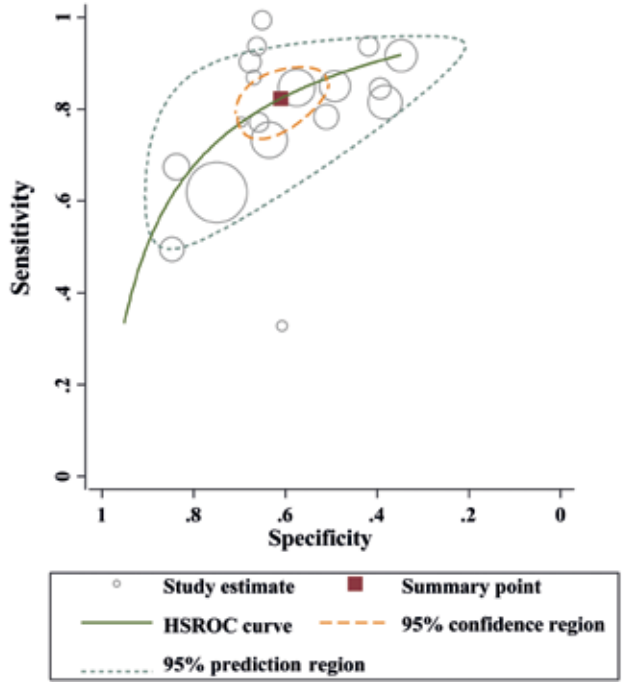


**Figure 3.** Hierarchical summary receiver-operating characteristic (HSROC) graph of the included studies that evaluated the diagnostic accuracy of FC for CRC detection

**Table 1.** Baseline characteristics of studies included in the meta-analysis

| Author (year) | Region | Patient spectrum | Disease type | Sample size | No. of CRC | No. of adenomas | Assay | Reference standard | Cut-off |
|---|---|---|---|---|---|---|---|---|---|
| Roseth et al. (13) 1993 | Norway | Known CRC, colorectal polyps | CRC, polyp | 206 | 53 | 40[a] | Roseth | Pathology, endoscopy | 10 mg/L |
| Kronborg et al. (14) 2000 | Europe | High-risk population for CRC | CRC, adenoma | 814 | 23 | 203 | PhiCal test | Pathology, endoscopy | 10 mg/L |
| Limburg et al. (15) 2000 | USA | Chronic diarrhea | Adenoma | 110 | - | 21 | PhiCal test | Pathology, endoscopy | 100 µg/g |
| Tibble et al. (16) 2000 | UK | Suspected inflammatory bowel disease | CRC, adenoma | 220 | 2 | 3 | Roseth | Pathology, endoscopy | 10 mg/L |
| Kristinsson et al. (17) 2001 | Norway | First-degree relatives of patients operated for CRC | CRC, adenoma | 237 | 9[b] | 73 | PhiCal test | Pathology, endoscopy | 10 mg/L |
| Tibble et al. (4) 2001 | UK | Suspected CRC | CRC, adenoma | 295 | 66 | 29 | Unclear | Pathology, endoscopy | 10 mg/L |
| Summerton et al. (18) 2002 | UK | Unselected patients with symptoms and/or signs | CRC, adenoma | 134 | 8 | 6 | PhiCal and Calprest ELISA | Pathology, endoscopy | 10 mg/L 50 µg/g |
| Tibble et al. (19) 2002 | UK | Suspected organic bowel disease | CRC | 346 | 7 | - | Roseth ELISA | Pathology, endoscopy | 10 mg/L |
| Carroccio et al. (20) 2003 | Italy | Chronic diarrhea | CRC | 80 | 3[c] | - | Calprest ELISA | Pathology, endoscopy | 50 µg/g |
| Costa et al. (21) 2003 | Italy | Outpatient clinic | CRC, adenoma | 239 | 18[d] | 8 | Calprest ELISA | Pathology, endoscopy, radiology | 50 µg/g |
| Limburg et al. (22) 2003 | USA | Suspected CRC | CRC, adenoma | 412 | 3 | 94 | PhiCal test | Pathology, endoscopy | 50 µg/g |
| Hoff et al. (23) 2004 | Norway | CRC screening in an average risk population | CRC, adenoma | 2,321 | 16 | 787 | PhiCal test | Pathology, endoscopy | 50 µg/g |
| Meucci et al. (24) 2010 | Italy | Unselected outpatients referred for colonoscopy | CRC, polyp | 870 | 34 | 244 | Calprest ELISA | Pathology, endoscopy | 50 mg/dL |
| Kok et al. (25) 2012 | Nether-lands | Suspected organic bowel disease | CRC, adenoma | 382 | 19 | 53 | Calprotectin POC test | Pathology, endoscopy | 50 µg/g |
| Parente et al. (26) 2012 | Italy | Suspected CRC | CRC, adenoma | 280 | 47 | 85 | Calprotectin Buhlmann ELISA | Pathology, endoscopy | 50 µg/g |
| Mowat et al. (1) 2016 | UK | Suspected CRC | CRC, high-risk adenoma | 755 | 28 | 41 | Calprotectin EK-CAL | Pathology, endoscopy | 50 µg/g |
| Rutka et al. (27) 2016 | Hungary | Suspected CRC | CRC, adenoma | 95 | 19 | 36 | Quantum Blue ELISA | Pathology, endoscopy | Unclear |

**Table 1.** Baseline characteristics of studies included in the meta-analysis (Continue)

| Author (year) | Region | Patient spectrum | Disease type | Sample size | No. of CRC | No. of adenomas | Assay | Reference standard | Cut-off |
|---|---|---|---|---|---|---|---|---|---|
| Turvill et al. (3) 2016 | UK | Suspected CRC | CRC, adenoma ≥ 10 mm | 654 | 39 | 33 | EK-CAL Calprotectin ELISA | Pathology, endoscopy, radiology | 50 µg/g |
| Högberg et al. (2) 2017 | Sweden | Suspected CRC | CRC, high-risk adenoma | 373 | 8 | 8 | CALPROV® Calprotectin ELISA Test | Pathology, endoscopy | 100 µg/g |
| Widlak et al. (28) 2017 | UK | Suspected CRC | CRC, adenoma | 430 | 25[b] | 42 | ELISA Calprotectin immunoassay | Pathology | 50 µg/g |

[a]Described as "polyps"; badenomas with high-grade dysplasia were considered as CRC; cdescribed as "colorectal cancer or adenomatous polyps"; dincluded two cases with colorectal lymphoma
CRC: colorectal cancer; ELISA: enzyme-linked immunosorbent assay

**Table 2.** Diagnostic performance of FC for colorectal neoplasia

| | CRC vs. non-CRC | Adenoma vs. non-adenoma |
|---|---|---|
| Number of patients | 8,913 | 7,662 |
| Sensitivity (95% CI) | 0.83 (0.77, 0.88) | 0.49 (0.37, 0.61) |
| I2 (%) | 61.72% | 92.79% |
| Specificity (95% CI) | 0.61 (0.54, 0.68) | 0.57 (0.49, 0.65) |
| I2 (%) | 97.98% | 97.84% |
| PLR (95% CI) | 2.15 (1.82, 2.55) | 1.14 (0.96, 1.35) |
| I2 (%) | 87.24% | 91.76% |
| NLR (95% CI) | 0.28 (0.21, 0.37) | 0.90 (0.76, 1.06) |
| I2 (%) | 64.99% | 84.97% |
| AUC (95% CI) | 0.81 (0.77, 0.84) | 0.55 (0.51, 0.59) |
| I2 (%) | 96% | 99% |
| DOR (95% CI) | 7.76 (5.41, 11.12) | 1.27 (0.91, 1.78) |
| I2 (%) | 97.97% | 100.00% |

FC: fecal calprotectin; CI: confidence interval; PLR: positive likelihood ratio; NLR: negative likelihood ratio; AUC: area under the HSROC curve; DOR: diagnostic odds ratio; HSROC: hierarchical summary receiver-operating characteristic

### *Diagnostic Accuracy of the FC test*

If multiple cut-off values or different assays for FC were reported in a single study, only the cut-off value or assay with the largest overall accuracy was used for the final analysis. Table 2 presents the results of the pooled sensitivity, specificity, AUC, and DOR of FC for distinguishing colorectal neoplasia. There were 18 studies that evaluated the diagnostic performance of FC for CRC detection (1-4, 13, 14, 16-21, 23-28). The pooled sensitivity and pooled specificity estimates were 0.83 (95% CI, 0.77-0.88) and 0.61 (95% CI, 0.54-0.68), respectively. Forest plots for pooled sensitivity and specificity are shown in Figure 2. The PLR and NLR values were 2.15 (95% CI, 1.82-2.55) and 0.28 (95% CI, 0.21-0.37), respectively, and the AUC and DOR values were 0.81 (95% CI, 0.77-0.84) and 7.76 (95% CI, 5.41-11.12), respectively (Figure 3).

Of the 20 eligible studies, 15 (1-4, 13-17, 21-23, 25, 26, 28) evaluated the diagnostic accuracy of FC for colorectal adenoma. For adenoma, the overall pooled sensitivity, specificity, PLR, and NLR for FC were 0.49 (95% CI, 0.37-0.61), 0.57 (95% CI, 0.49-0.65), 1.14 (95% CI, 0.96-1.35), and 0.90 (95% CI, 0.76-1.06), respectively. The average AUC and DOR values for adenoma were 0.55 (95% CI, 0.51-0.59) and 1.27 (95% CI, 0.91-1.78), respectively. Overall, these results suggested that the utility of FC for detecting colorectal adenoma was insufficient for clinical application (Table 2).

### *Publication Bias*

Deek's funnel plot asymmetry and the results of the Deek's test (p=0.204 for CRC, p=0.407 for adenoma) did not suggest any potential publication bias (Figure 4).

### *Investigation of Heterogeneity*

Our results were heterogeneous (Table 2). We first performed a Spearman analysis and found that there was no threshold effect within the studies (r=0.451, p=0.060). We also performed subgroup analyses to detect the source of heterogeneity based on patient spectrum, FC assay method, prevalence, sample size, publication year, and cut-off value. Synthesized data revealed that the heterogeneity did not significantly change with variations in the patient spectrum (Table 3). CRC detection by the conventional FC assay exhibited greater sensitivity than the new assay, as evidenced by the reduced heterogeneity ($I^2$=37.42%) and an increased DOR value of 12.15 (5.67, 26.05). Therefore, the assay by which FC was mea-

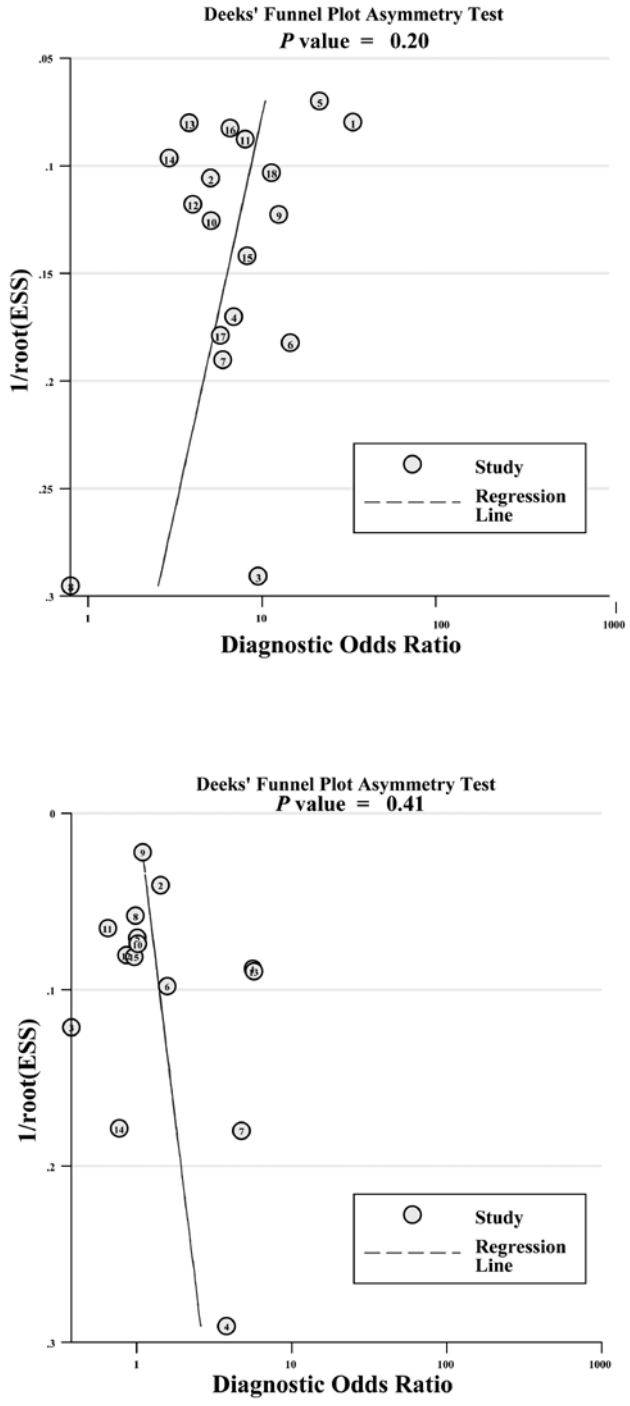


**Figure 4.** Funnel plot with superimposed regression line for studies of (A) CRC and (B) colorectal adenoma

sured appeared to significantly contribute to the heterogeneity. Probably, the performance characteristics of FC for CRC detection were considerably better in studies with a higher level of prevalence (above the average) and a smaller sample size (below the median). However, the diagnostic performance remained similar and the heterogeneity did not notably decrease with respect to the publication year or cut-off value (Table 3). We also performed meta-regression to identify any significant sources of heterogeneity (Figure S2). The patient spectrum ($p<0.01$), FC assay ($p<0.001$), sample size ($p<0.001$), and publication year ($p<0.001$) significantly contributed to the heterogeneity in sensitivity. However, the cut-off value ($p<0.05$) accounted for heterogeneity in specificity.

### *Predictive Value of the FC Test*

The Fagan plot was applied to assess the clinical utility of the FC test for predicting CRC in the screened populations (Figure S3). The diagnosis of CRC was confirmed in 4.6% (n=427) of the total included patients (n=9,253). Consequently, with a pretest probability of 4.6%, a positive FC test was found to predict an increase in the risk of CRC to 10%; conversely, a negative FC test was found to decrease the risk of CRC to 1%.

## *DISCUSSION*

In this study, we conducted a robust systematic review and meta-analysis. The results showed that the pooled sensitivity and specificity values of FC for CRC were approximately 83% and 61%, respectively. In addition, the overall accuracy of FC for predicting CRC was 81%. With regard to colorectal adenoma, the FC test is of inferior diagnostic value as assessed by our quantitative analysis (AUC, 0.55; 95% CI, 0.51-0.59; DOR, 1.27; 95% CI, 0.91-1.78).

The latest systematic review was published in 2007, which included 12 studies with respect to the diagnostic accuracy of FC for colorectal neoplasia (6). Our present analysis included 20 studies and it updates the body of evidence. The prior systematic review combined CRC and adenoma together as "neoplasia" and concluded that FC was not suitable for CRC screening. Our study differs in that we analyzed each case of CRC and adenoma. Nevertheless, our results suggest that the FC test, as it currently exists, cannot be recommended for CRC detection.

The diagnosis of CRC can be challenging for gastroenterologists because of its nonspecific symptoms. At present, the majority of referred cases from primary care for evaluation of bowel symptoms undergo invasive tests

**Table 3.** Subgroup analyses of FC for CRC detection

| Subgroup | No. of studies | Sensitivity (95% CI) | Specificity (95% CI) | PLR (95% CI) | NLR (95% CI) | AUC (95% CI) | DOR (95% CI) |
|---|---|---|---|---|---|---|---|
| *Patient spectrum* | | | | | | | |
| Suspected CRC | 9 | 0.81 (0.73, 0.88) | 0.63 (0.49, 0.75) | 2.19 (1.63, 2.96) | 0.30 (0.21, 0.41) | 0.81 (0.78, 0.84) | 7.42 (4.66, 11.82) |
| $I^2$ (%) | | 69.72% | 98.61% | 93.16% | 69.54% | 96% | 99.60% |
| Others | 8 | 0.85 (0.73, 0.92) | 0.59 (0.52, 0.66) | 2.09 (1.76, 2.48) | 0.25 (0.14, 0.45) | 0.76 (0.72, 0.80) | 8.28 (4.21, 16.26) |
| $I^2$ (%) | | 62.87% | 97.76% | 49.09% | 69.17% | 80% | 96.00% |
| *FC assay* | | | | | | | |
| Old | 6 | 0.88 (0.78, 0.93) | 0.63 (0.58, 0.68) | 2.39 (1.98, 2.87) | 0.20 (0.11, 0.36) | 0.80 (0.77, 0.84) | 12.15 (5.67, 26.05) |
| $I^2$ (%) | | 37.42% | 90.38% | 10.33% | 41.39% | 0% | 90.99% |
| *New* | 9 | 0.78 (0.68, 0.86) | 0.59 (0.45, 0.72) | 1.92 (1.47, 2.50) | 0.36 (0.27, 0.49) | 0.77 (0.73, 0.81) | 5.30 (3.55, 7.93) |
| $I^2$ (%) | | 76.48% | 99.00% | 90.80% | 76.68% | 97.00% | 96.97% |
| *Prevalence* | | | | | | | |
| >average | 9 | 0.87 (0.81, 0.92) | 0.59 (0.48, 0.70) | 2.14 (1.66, 2.77) | 0.21 (0.15, 0.31) | 0.85 (0.82, 0.88) | 10.07 (6.04, 16.77) |
| $I^2$ (%) | | 58.07% | 97.59% | 90.89% | 52.41% | 93% | 98.93% |
| <average | 9 | 0.75 (0.64, 0.83) | 0.64 (0.54, 0.72) | 2.07 (1.68, 2.53) | 0.40 (0.29, 0.55) | 0.76 (0.72, 0.79) | 5.22 (3.41, 7.99) |
| $I^2$ (%) | | 40.98% | 98.26% | 61.72% | 48.60% | 92% | 83.80% |
| *Sample size* | | | | | | | |
| >median | 9 | 0.78 (0.68, 0.86) | 0.62 (0.49, 0.73) | 2.06 (1.59, 2.67) | 0.35 (0.26, 0.47) | 0.78 (0.74, 0.82) | 5.87 (3.98, 8.65) |
| $I^2$ (%) | | 61.02% | 98.93% | 89.60% | 45.95% | 96% | 91.34% |
| <median | 9 | 0.88 (0.84, 0.92) | 0.61 (0.53, 0.68) | 2.25 (1.85, 2.74) | 0.19 (0.13, 0.28) | 0.89 (0.86, 0.91) | 11.86 (6.87, 20.47) |
| $I^2$ (%) | | 52.74% | 90.99% | 78.89% | 71.70% | 35% | 99.71% |
| *Publication year* | | | | | | | |
| Before 2007 | 10 | 0.86 (0.76, 0.92) | 0.63 (0.57, 0.69) | 2.33 (2.00, 2.70) | 0.22 (0.13, 0.39) | 0.78 (0.74, 0.82) | 10.49 (5.59, 19.72) |
| $I^2$ (%) | | 60.69% | 96.25% | 54.46% | 68.24% | 83% | 93.57% |
| After 2007 | 8 | 0.80 (0.71, 0.87) | 0.60 (0.44, 0.73) | 1.99 (1.47, 2.69) | 0.33 (0.25, 0.44) | 0.79 (0.76, 0.83) | 6.00 (3.88, 9.28) |
| $I^2$ (%) | | 66.65% | 98.64% | 91.38% | 60.78% | 96% | 97.75% |
| Cut-off value | | | | | | | |
| 50 µg/g or 10 mg/L | 15 | 0.84 (0.78, 0.89) | 0.59 (0.51, 0.67) | 2.06 (1.73, 2.45) | 0.27 (0.19, 0.37) | 0.81 (0.77, 0.84) | 7.71 (5.06, 11.76) |
| $I^2$ (%) | | 60.61% | 98.12% | 87.29% | 65.00% | 95% | 97.56% |
| Others | 3 | - | - | - | - | - | - |
| $I^2$ (%) | | - | - | - | - | - | - |

AUC, area under the HSROC curve; FC, fecal calprotectin; CI, confidence interval; CRC, colorectal cancer; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR diagnostic odds ratio; HSROC, Hierarchical summary receiver-operating characteristic

including colonoscopy, but the diagnosis of pathology is low (1). For example, currently, only 8% of patients who complain of bowel symptoms receive a confirmed diagnosis of CRC (3). Also, functional gastrointestinal disorders, including irritable bowel syndrome, are often found to coexist in 75% of symptomatic patients (as assessed with colonoscopy), and such conditions respond well to the typical treatment (3,26). However, colonoscopy is an invasive, expensive procedure with an established risk of complications. Moreover, there is limited availability of colonoscopy services in many regions of the world; thus, colonoscopy screening should be reserved for higher risk individuals to optimize resource utilization (26). Therefore, a non-invasive and simple test may be preferable for prioritizing referrals for colonoscopy. The fecal occult blood test has been widely used as a preliminary assessment, and it has been shown to reduce disease-specific mortality; however, its accuracy is affected by the dietary pattern, and nearly 50% of colonoscopies performed based on a positive fecal occult blood test result show no evidence of neoplasia (26). Indeed, previous studies have supported the use of the fecal immunohistochemical test (FIT) for CRC screening in average-risk populations because of its high accuracy and reproducibility (29). FC is stable in stools because of its resistance to bacterial degradation at room temperature for up to 7 days, and studies have found that the sensitivity of the FC test for detecting CRC ranges from 79% to 95% (2). However, our data suggest that the FC test alone is not powerful enough to screen for CRC. Therefore, we infer that there is probably no place for FC as a screening tool for CRC in average-risk populations.

Although the FC test cannot be recommended as the sole screening modality for CRC, our data still have clinical implications. We propose that FC should be used as an auxiliary tool to help clinicians stratify CRC risk in different patients because the FC test may help predict CRC development. Nevertheless, significant risk factors associated with CRC, such as age, gender, anemia, family history of cancer, and change of bowel movement habits, should also be considered when planning further treatment. Indeed, previously reported preliminary results substantiated the benefit of the FC test in a diagnostic model of significant colorectal disease (including CRC) (30). It is thus imperative that future studies include larger and better-defined patient populations.

There was significant heterogeneity among some of our results. We selected studies with a prospective design and excluded case-control studies because they tend to overestimate the diagnostic precision. Therefore, the heterogeneity caused by the study design was largely avoided. We used a random-effects model, which assumed that the true effects were normally distributed, and more weightage was assigned to the small-sized studies compared with the fixed-effects model. Subgroup analyses were also conducted to address heterogeneity (Table 3). However, the subgroup summary estimates did not differ significantly from the overall summary estimates, suggesting that the overall summary measures reflect the reasonable estimations of the overall FC test accuracy despite some statistical heterogeneity. Selection bias was likely to be another source of heterogeneity because numerous different diagnoses were grouped together in some of our analyses. For instance, the non-CRC group or non-adenoma group contained patients with diverse gastrointestinal conditions, such as inflammatory bowel disease, diverticulosis, food intolerance, and irritable bowel syndrome, as well as healthy controls. Nonetheless, this phenomenon reflects a "real world" scenario and therefore provides a realistic estimation of the CRC risk. Moreover, meta-regression was used to decipher the nature of heterogeneity. Patient spectrum, FC assay method, sample size, publication year, and cut-off value were identified as potential confounding factors that may have contributed to the heterogeneity.

Some limitations of our study merit consideration. First, the FC assay varied across studies, and this affected our interpretation of the analyses. Second, we could not determine the sensitivity and specificity of FC for CRC stratified by disease site (proximal or distal) or cancer stage because of a paucity of relevant data. Third, few studies focused on the value of simultaneously or consecutively administering patients with other tests. For example, Mowat *et al.* (1) found FIT to be superior to FC, whereas others found a similar negative predictive value when comparing FC and FIT (25,26). Overall, three studies assessed the value of the combination of FC and FIT for detecting CRC but showed no significant improvement in the diagnostic accuracy (1,2,28). However, the resolution of this issue will require more data and additional studies. These approaches could aid clinicians to maximize the use of FC and other biomarkers in clinical practice. Fourth, our study could not derive a summarized estimation of the performance of FC for advanced adenoma because the definition of advanced adenoma varied across studies. Finally, the absence of data from Asia may compromise the results. Several of the included studies were conducted in primary care facilities. However, in Asia, particularly mainland China, the public prefers to select hospital-based services instead of com-

munity-based primary care against a background of the "barefoot doctor" legacy in rural China and hence, a mistrust of doctors. China has the largest population in the world; however, cancer surveillance needs improvement, and the availability of endoscopy in secondary-care facilities lags far behind other areas of the world. Therefore, studies are needed to address this deficiency.

In summary, the FC test was found to be unsuitable for screening CRC. Whether the combination of FC and other biomarkers would enhance CRC detection needs further investigation. Moreover, studies from Asia are urgently needed.

You can reach the supplemental figures of this article at https://10.5152/tjg.2018.17606

***Ethics Committee Approval:*** Ethics committee approval was received for this study from the Ethics Committee of Zhejiang University Jinhua Hospital.

***Informed Consent:*** N/A.

***Peer-review:*** Externally peer-reviewed.

***Author Contributions:*** Concept - X.Y.; Design - J.D.; Supervision - J.D.; Resource - J.D.; Materials - J.H.; Data Collection and/or Processing - J.H.; Analysis and/or Interpretation - X.Y.; Literature Search - X.Y.; Writing - X.Y. Critical Reviews - J.D.

***Conflict of Interest:*** The authors have no conflict of interest to declare.

***Financial Disclosure:*** This study was funded by two grants from Key Research Projects of Jinhua Municipal Technology Bureau (Grant No. 2013-3-010) and the Science and Technology Bureau of Jinhua, Zhejiang Province, China (Grant No. 2014-3-053).

## REFERENCES

1. Mowat C, Digby J, Strachan JA, et al. Faecal haemoglobin and faecal calprotectin as indicators of bowel disease in patients presenting to primary care with bowel symptoms. Gut 2016; 65: 1463-9. [CrossRef]
2. Hogberg C, Karling P, Rutegard J, Lilja M. Diagnosing colorectal cancer and inflammatory bowel disease in primary care: The usefulness of tests for faecal haemoglobin, faecal calprotectin, anaemia and iron deficiency. A prospective study. Scand J Gastroenterol 2017; 52: 69-75. [CrossRef]
3. Turvill J, Aghahoseini A, Sivarajasingham N, et al. Faecal calprotectin in patients with suspected colorectal cancer: a diagnostic accuracy study. Br J Gen Pract 2016; 66: e499-506. [CrossRef]
4. Tibble J, Sigthorsson G, Foster R, Sherwood R, Fagerhol M, Bjarnason I. Faecal calprotectin and faecal occult blood tests in the diagnosis of colorectal carcinoma and adenoma. Gut 2001; 49: 402-8. [CrossRef]
5. Lehmann FS, Trapani F, Fueglistaler I, et al. Clinical and histopathological correlations of fecal calprotectin release in colorectal carcinoma. World J Gastroenterol 2014; 20: 4994-9. [CrossRef]
6. von Roon AC, Karamountzos L, Purkayastha S, et al. Diagnostic precision of fecal calprotectin for inflammatory bowel disease and colorectal malignancy. Am J Gastroenterol 2007; 102: 803-13. [CrossRef]
7. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working G. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008; 149: 889-97. [CrossRef]
8. Lin JF, Chen JM, Zuo JH, et al. Meta-analysis: fecal calprotectin for assessment of inflammatory bowel disease activity. Inflamm Bowel Dis 2014; 20: 1407-15. [CrossRef]
9. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529-36. [CrossRef]
10. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58: 982-90. [CrossRef]
11. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001; 20: 2865-84. [CrossRef]
12. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005; 58: 882-93. [CrossRef]
13. Roseth AG, Kristinsson J, Fagerhol MK, et al. Faecal calprotectin: a novel test for the diagnosis of colorectal cancer? Scand J Gastroenterol 1993; 28: 1073-6. [CrossRef]
14. Kronborg O, Ugstad M, Fuglerud P, et al. Faecal calprotectin levels in a high risk population for colorectal neoplasia. Gut 2000; 46: 795-800. [CrossRef]
15. Limburg PJ, Ahlquist DA, Sandborn WJ, et al. Fecal calprotectin levels predict colorectal inflammation among patients with chronic diarrhea referred for colonoscopy. Am J Gastroenterol 2000; 95: 2831-7. [CrossRef]
16. Tibble J, Teahon K, Thjodleifsson B, et al. A simple method for assessing intestinal inflammation in Crohn's disease. Gut 2000; 47: 506-13. [CrossRef]
17. Kristinsson J, Nygaard K, Aadland E, et al. Screening of first degree relatives of patients operated for colorectal cancer: evaluation of fecal calprotectin vs. hemoccult II. Digestion 2001; 64: 104-10. [CrossRef]
18. Summerton CB, Longlands MG, Wiener K, Shreeve DR. Faecal calprotectin: a marker of inflammation throughout the intestinal tract. Eur J Gastroenterol Hepatol 2002; 14: 841-5. [CrossRef]
19. Tibble JA, Sigthorsson G, Foster R, Forgacs I, Bjarnason I. Use of surrogate markers of inflammation and Rome criteria to distinguish organic from nonorganic intestinal disease. Gastroenterology 2002; 123: 450-60. [CrossRef]
20. Carroccio A, Iacono G, Cottone M, et al. Diagnostic accuracy of fecal calprotectin assay in distinguishing organic causes of chronic diarrhea from irritable bowel syndrome: a prospective study in adults and children. Clin Chem 2003; 49: 861-7. [CrossRef]

21. Costa F, Mumolo MG, Bellini M, et al. Role of faecal calprotectin as non-invasive marker of intestinal inflammation. Dig Liver Dis 2003; 35: 642-7. [CrossRef]
22. Limburg PJ, Devens ME, Harrington JJ, Diehl NN, Mahoney DW, Ahlquist DA. Prospective evaluation of fecal calprotectin as a screening biomarker for colorectal neoplasia. Am J Gastroenterol 2003; 98: 2299-305. [CrossRef]
23. Hoff G, Grotmol T, Thiis-Evensen E, Bretthauer M, Gondal G, Vatn MH. Testing for faecal calprotectin (PhiCal) in the Norwegian Colorectal Cancer Prevention trial on flexible sigmoidoscopy screening: comparison with an immunochemical test for occult blood (FlexSure OBT). Gut 2004; 53: 1329-33. [CrossRef]
24. Meucci G, D'Inca R, Maieron R, et al. Diagnostic value of faecal calprotectin in unselected outpatients referred for colonoscopy: A multicenter prospective study. Dig Liver Dis 2010; 42: 191-5. [CrossRef]
25. Kok L, Elias SG, Witteman BJ, et al. Diagnostic accuracy of point-of-care fecal calprotectin and immunochemical occult blood tests for diagnosis of organic bowel disease in primary care: the Cost-Effectiveness of a Decision Rule for Abdominal Complaints in Primary Care (CEDAR) study. Clin Chem 2012; 58: 989-98. [CrossRef]
26. Parente F, Marino B, Ilardo A, et al. A combination of faecal tests for the detection of colon cancer: a new strategy for an appropriate selection of referrals to colonoscopy? A prospective multicentre Italian study. Eur J Gastroenterol Hepatol 2012; 24: 1145-52. [CrossRef]
27. Rutka M, Bor R, Balint A, et al. Diagnostic Accuracy of Five Different Fecal Markers for the Detection of Precancerous and Cancerous Lesions of the Colorectum. Mediators Inflamm 2016; 2016: 2492081. [CrossRef]
28. Widlak MM, Thomas CL, Thomas MG, et al. Diagnostic accuracy of faecal biomarkers in detecting colorectal cancer and adenoma in symptomatic patients. Aliment Pharmacol Ther 2017; 45: 354-63. [CrossRef]
29. Katsoula A, Paschos P, Haidich AB, Tsapas A, Giouleme O. Diagnostic Accuracy of Fecal Immunochemical Test in Patients at Increased Risk for Colorectal Cancer: A Meta-analysis. JAMA Intern Med 2017; 177: 1110-8. [CrossRef]
30. Elias SG, Kok L, de Wit NJ, et al. Is there an added value of faecal calprotectin and haemoglobin in the diagnostic work-up for primary care patients suspected of significant colorectal disease? A cross-sectional diagnostic study. BMC Med 2016; 14: 141. [CrossRef]

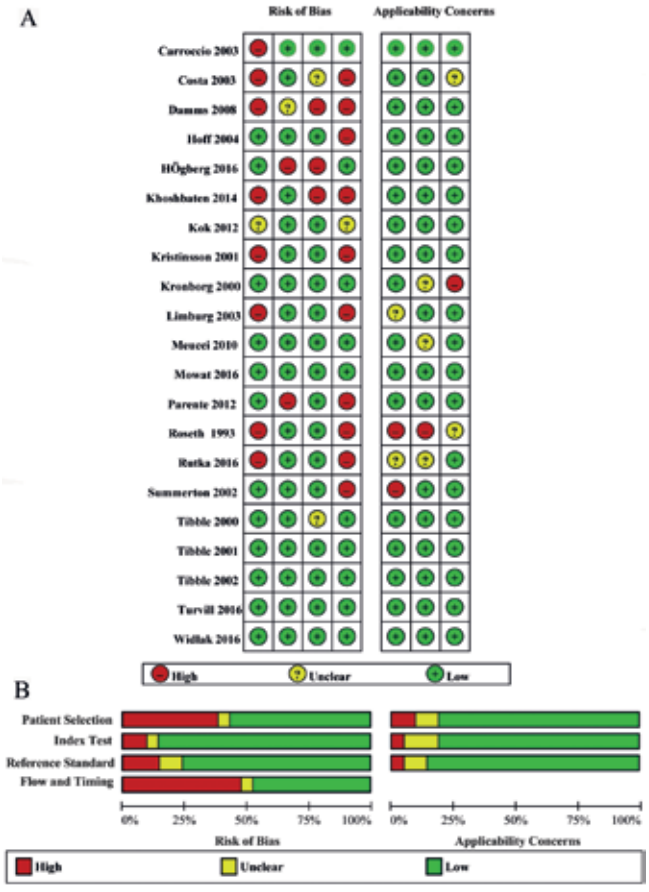


**Figure S1.** Risk of bias and applicability concerns: (A) Summary information (B) Graphical depiction

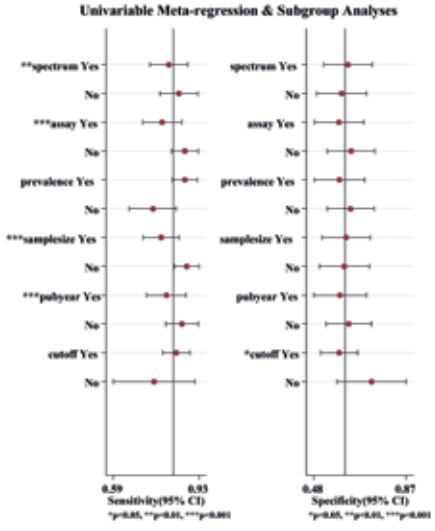


**Figure S2.** Meta-regression and subgroup analysis for sources of heterogeneity in CRC

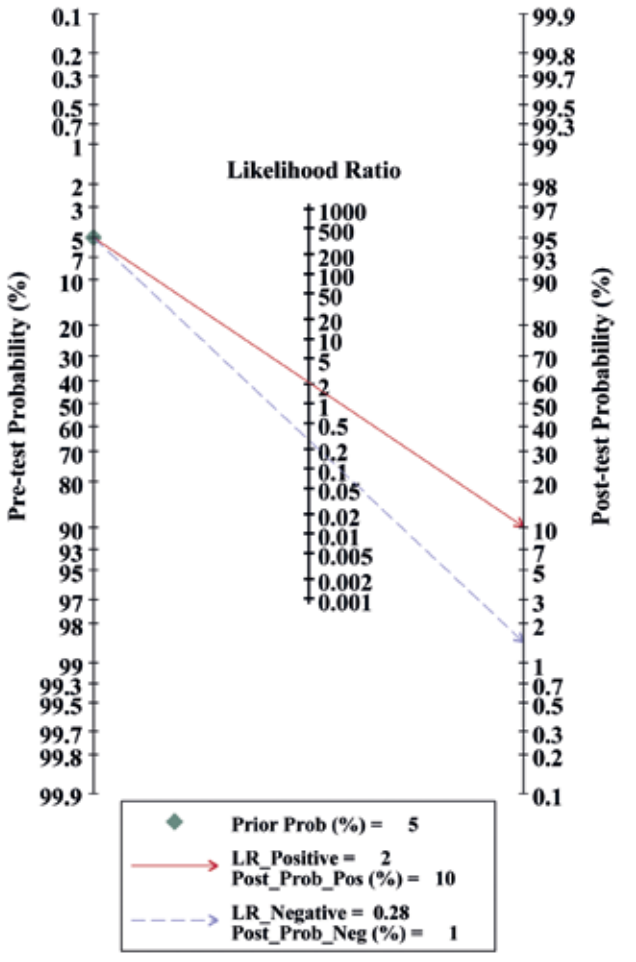


**Figure S3.** Fagan's plot for the post-test probability of CRC after an FC-positive result (upper line) or FC-negative result (lower line)