Test Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-Analysis

Thomas F. Imperiale, MD[1-3], Rachel N. Gruber, MS[2], Timothy E. Stump, MA[4], Thomas Emmett, MD,[5] Patrick O. Monahan, PhD[4].

From the Departments of Medicine[1] and Biostatistics[4], Indiana University School of Medicine; the Center for Innovation, Richard L. Roudebush VA Medical Center[2]; and Regenstrief Institute, Inc.[3], and the Ruth Lilly Medical Library[5], Indianapolis, IN.

Running title: Test Characteristics of Fecal Immunochemical Tests

Key words: colorectal cancer screening; colorectal neoplasia; fecal immunochemical test; systematic review

Word count (abstract): 358

Word count (manuscript): 3771

Number of tables: 1

Number of figures: 3

Number of appendix tables / figures: 10

Corresponding author:

Thomas F. Imperiale, MD
Indiana University Medical Center
Regenstrief Institute
1100 West 10th Street
Indianapolis, IN 40202
TEL: 317-274-9046
FAX: 317-274-9304
E-mail: timperia@iu.edu

This is the author's manuscript of the article published in final edited form as:

Imperiale, T. F., Gruber, R. N., Stump, T. E., Emmett, T. W., & Monahan, P. O. (2019). Performance Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis. Annals of Internal Medicine, 170(5), 319–329. https://doi.org/10.7326/M18-2390

ABSTRACT

Background: Studies report inconsistent performance characteristics of fecal immunochemical tests (FITs) for both colorectal cancer (CRC) and advanced adenomas.

Purpose: To summarize test characteristics of fecal immunochemical tests (FITs) for CRC and advanced adenomas in average-risk persons undergoing screening colonoscopy (reference standard) and identify factors affecting these characteristics.

Data Sources: Ovid MEDLINE, PubMed, Embase, and the Cochrane Library from database inception through October 2018; reference lists of studies and reviews.

Study Selection: Two reviewers independently screened records to identify published English-language prospective or retrospective observational studies that evaluated FIT sensitivity and specificity for colonoscopy findings in asymptomatic average-risk adults.

Data Extraction: Two authors independently extracted study data and evaluated study quality.

Data Synthesis: We included 31 studies (120,255 participants, 17 FITs), all of which we judged had low-to-moderate risk for bias. Performance characteristics depended on the threshold for a positive test. A test threshold of 10 µg/g feces resulted in a CRC sensitivity of 0.91 (95% CI, 0.84 to 0.95) and negative likelihood ratio of 0.10 (CI, 0.06 to 0.19), while a threshold of > 20 µg/g resulted in CRC specificity of 0.95 (CI, 0.94 to 0.96) and positive likelihood ratio of 15.49 (CI, 9.82 to 22.39). Advanced adenoma sensitivity was 0.40 (CI, 0.33 to 0.47) and negative likelihood ratio was 0.67 (CI, 0.57 to 0.78) at 10 µg/g, while specificity was 0.95 (0.94 to 0.96) and positive likelihood ratio was 5.86 (CI, 3.77-8.97) at > 20 µg/g. There was low-to high heterogeneity among studies, depending on threshold. While several FITs had adequate test performance, CRC sensitivity and specificity for one qualitative FIT were 90% and 91%, respectively, at its single threshold of 10 µg/g, with positive and negative likelihood ratios of 10.13 and 0.11, respectively. Comparison of performance of 3 FITs at 3 thresholds was inconclusive: CIs overlapped and the comparisons were across, rather than within, studies.

Limitations: Only English-language studies were included; incomplete reporting limited quality assessment of some evidence. Test characteristics are for one-time, rather than for serial, testing.

Conclusion: Single-application FITs have moderate-to-high CRC sensitivity and specificity depending on the positivity threshold. Sensitivity of one-time testing for advanced adenomas is low regardless of threshold.

Primary Funding Source: Department of Medicine, Indiana University School of Medicine.

BACKGROUND

Colorectal cancer (CRC) is a leading cause of death among digestive diseases and the second leading cause of cancer related death in the United States (1). Despite the effectiveness and cost-effectiveness of screening (2-4), just 60-65% of the screen-eligible population is current with screening (5), a rate that has fallen short of the targeted goal of 80% by 2018 (2, 5, 6). To some extent, this shortcoming represents concern over the best test and strategy for screening. Although screening colonoscopy is the most frequently used test in the U.S. (5), several other countries screen with annual or biennial stool blood tests or a combination of stool testing and lower endoscopy (7, 8).

While studies show guaiac-based fecal occult blood testing (gFOBT) reduces CRC incidence and mortality (9-13), gFOBT has several shortcomings. Limitations include low single-application sensitivity for CRC, poor detection of advanced adenomas, the need for dietary and medication restrictions, and requirement of more than one specimen. Fecal immunochemical testing (FIT) of stool for human globulin is more sensitive and specific than gFOBT for colorectal cancer and advanced adenomas, and has higher rates of participation and acceptance (14-16). Studies evaluating the test characteristics of FIT, however, show inconsistent findings for CRC and advanced adenomas, the latter of which includes adenomas $\geq$ 1 cm and those with villous histology and/or high-grade dysplasia. A systematic review published in 2014 summarized FIT test performance for CRC (17), but did not quantify test characteristics for advanced adenomas. The objectives of this systematic review and meta-analysis are to: update the summarization of FIT performance for CRC; quantify FIT test characteristics for advanced adenomas; and evaluate whether variation in reported test characteristics among studies is a function of the threshold used to define a positive test or of the specific test brand.

METHODS

Rather than develop and register a new formal protocol, we used two prior systematic reviews as guides for our study methodology (17, 18). We followed standard procedures for systematic reviews and reported results according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (19, 20).PLEASE USE THE 2009 ANNALS PRISMA STATEMENT article for this reference.

Data Sources and Searches

We did English-language literature searches of the MEDLINE, EMBASE, and Cochrane databases from inception to 17 October 2018 to identify studies assessing test performance of one or more FITs. Searches were done using various combinations of the following terms: feces, occult blood, colon cancer, cancer screening, early diagnosis, immunochemistry, and FIT (See Appendix Table 1). We also reviewed reference lists of relevant systematic reviews and meta-analyses (17, 18) and of articles that met selection criteria.

Study Selection

Two authors independently screened all titles and potentially relevant abstracts, and then full texts of articles that we thought were potentially eligible. Inclusion criteria were published English-language prospective or retrospective observational studies that evaluated FIT sensitivity and specificity in asymptomatic average-risk adults and that used colonoscopy as the reference standard. Data available only in abstract form and grey literature were not eligible.

Data Extraction and Quality Assessment

Two reviewers reviewed descriptive and quantitative data from each study. Data extraction was done primarily by author RG and independently validated by authors TI or TE. For each study, we extracted data on sample size, mean age, brand(s) of FIT used, thresholds for positivity

(expressed as micrograms [μg] of hemoglobin per gram [g] of feces), numbers of participants with CRC and AA, and test characteristics for CRC and advanced adenomas. When available, raw data on CRC and advanced adenomas were extracted. When only computed data were available, individual raw data were calculated based on identified proportions. When data were missing from articles, the corresponding authors were contacted. When more than one FIT cutoff or threshold was used, test characteristics for thresholds commonly used were extracted. Two authors (among TI, RG, and TE) independently assessed study characteristics and evaluated study quality by using the revised version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 tool (21). Discrepancies between reviewers for study quality assessment were resolved by discussion.

Data Synthesis and Analysis

Sensitivity and specificity at one or more threshold were reported for each study. We combined studies and report results for both CRC and advanced adenomas, based on test threshold, in micrograms of hemoglobin per gram of feces, where we grouped studies with FIT thresholds of < 10 μg/g, 10 μg/g, > 10 to < 20 μg/g, 20 μg/g, and > 20 μg/g. To assess statistical heterogeneity, we quantified the I^2 measure, which indicates the percentage of total variation across studies due to heterogeneity rather than chance (22). For all summary-level estimates, we used a bivariate generalized linear mixed model to simultaneously estimate pooled measures of sensitivity and specificity separately for both CRC and advanced adenomas while accounting for the potential correlation between sensitivity and specificity. The bivariate approach produces unbiased estimates of sensitivity, specificity and their correlation (23) and does not rely on an ad hoc continuity correction for zero marginal counts. Likelihood ratios (LR) were calculated using the bivariate estimates as follows: [LR+ = sensitivity / (1 – specificity); LR- = (1 – sensitivity) / specificity]. Summary receiver-operating characteristic curves (SROC) were obtained along with 95% confidence regions for the bivariate estimates of sensitivity and 1-

specificity. We also combined studies by brand of FIT and did so by threshold to enable indirect comparisons.

Meta-Disc software (Hospital Universitario Ramón y Cajal) (24) was used to provide the I^2 measure. For all other summary estimates of test characteristics, the glmer function (25) of the lme4 package (26) for R (R Foundation for Statistical Computing) (27) was used to estimate the bivariate generalized mixed models.

Role of the Funding Source

The study was funded by the Department of Medicine and the Melvin & Bren Simon Cancer Center at the Indiana University School of Medicine and by the Regenstrief Institute, Inc. These funding sources had no role in the conception, design, conduct, analysis, or review of the manuscript or the decision to submit it for publication.

RESULTS

Our search strategy (Appendix Table 1) generated 4976 citations, of which 31 articles were included in the analysis (Figure 1) (Appendix Table 2) (28-58). Studies were published between 2001 and 2018. Funding sources included federal government (n=15), private or intramural (n=4), and corporate (n=2) sources, and were not reported for 10 studies. Participant sample size from all studies totaled 120,255 and ranged from 284 to 21,805. Mean age ranged from 48.2 years to 64 years. All study populations were composed of asymptomatic, mostly average-risk persons in the screening age range (generally 50 to 75 years old) who enrolled in cancer prevention programs of screening colonoscopy. Persons with prior colorectal neoplasia, inflammatory bowel disease, high-risk family history, or colonoscopy within the previous 5-10 years were excluded, as were (post-hoc) those in whom bowel prep quality was unacceptable or

extent of colonoscopy incomplete.

Eighteen different FITs were tested, ranging from one to six FITs tested in a single study. OC Sensor (Eiken Chemical Co., LTD) was tested in 14 (58%) studies (35, 39-42, 46-49, 51, 52, 54, 56, 57), including OC-FIT-CHEK (Eiken Chemical Co., LTD) in two (48, 57) of those 14 studies, OC Light (Eiken Chemical Co., LTD) in 5 studies (29, 36, 37, 43, 45), and OC Hemodia (Eiken Chemical Co., LTD) and FOB-Gold (Sentinel Diagnostics) in 3 studies each (28, 30, 32, 33, 55, 58). Many of the remaining FITs are or were available only within a single country or region. Thresholds for positivity ranged from 2 μg hemoglobin per gram (g) of feces to 67 μg/g feces, with 10 studies using a positivity threshold of < 10 μg/g, 16 studies using a threshold of 10 μg/g, 8 studies using a threshold of 11-19 μg/g, and 26 studies using a threshold of ≥ 20 μg/g.

FIT test characteristics for CRC and advanced adenomas based on each threshold tested for individual studies are shown in Appendix Table 2. All studies assessed the sensitivity and specificity of one or more FITs for advanced adenoma, which ranged in prevalence from 1.26% to 12.2%, while all but three studies (33, 39, 51) did so for CRC, which ranged in prevalence from 0.15% to 3.48%.

We judged the quality of most studies as high (Appendix Figure 1). All were cross-sectional. Only one used a prospective case-control design (46); for this study, we utilized data only from the control group, which, like the other studies, was composed of persons undergoing screening colonoscopy. A lack of detail in study methods precluded knowing whether a consecutive or random sample of persons participated for nearly half of the studies. Despite this, we assessed most studies as having low-to moderate risk for selection bias. For approximately 30% of the studies it was unclear whether FIT results and colonoscopic findings were interpreted independently of one another. For 25% of the studies, the interval between FIT and

colonoscopy was not specified.

The studies tested one or more FITs; several studies used more than one threshold with the objective of determining the optimal threshold. FIT was collected prior to colonoscopy in all but one study,(49) which provided no information. While five studies did not specifically state that the FITs were interpreted without knowledge of the colonoscopic findings (28, 33, 39, 49, 56), FIT processing was automated, making it unlikely that FIT interpretation was biased by the findings. Colonoscopy was the reference standard in all studies. In 20 (65%) of 31 articles (29, 31, 33-38, 41-48, 50, 51, 54, 58), authors reported that colonoscopy was performed blinded to FIT results, while authors of 11 articles (28, 30, 32, 39, 40, 49, 52, 53, 55-57) made no comment about blinding. Overall, we assessed the risk for bias in interpretation of colonoscopic findings due to FIT results to be low. Several studies had post-hoc exclusions of subjects due to not completing colonoscopy or FIT, unsatisfactory quality of bowel preparation, incomplete colonoscopy, or other reasons. (Appendix Table 3). Only 3 studies provided information on indeterminate FITs and for only one of three studies was it clear that test results were truly indeterminate (Appendix Table 3). The number of participants excluded due to an unsatisfactory colonoscopy was provided in nearly half of the studies and ranged from 0.12% to 8.9%.The risk for biased patient flow affecting validity of FIT test characteristics was assessed to be low (Appendix Figure 1).

Overall test characteristics of FIT

I^2 values for heterogeneity were in the low-to-high range for sensitivity and were high for specificity for all FIT thresholds for both CRC and AA (Appendix Table 4). Figures 2 and 3 show the main results for FIT sensitivity, specificity, and likelihood ratios for CRC and advanced adenoma, respectively. Among 31 studies that included a total of 58 assessments of several FITs at various thresholds, CRC sensitivity ranged from 91% (CI, 84 to 95%) for a threshold of

10 µg/g feces to 71% (CI, 56 to 83%) for a threshold of > 20 µg/g, while specificity ranged from 90% (CI, 81 to 95%) for a threshold of < 10 µg/g feces to 95% (CI, 94 to 96%) for a threshold of ≥ 20 µg/g (Figure 2). Corresponding likelihood ratios ranged from a positive likelihood ratio of 15.49 at a threshold of > 20 µg/g to a negative likelihood ratio of 0.10 at a threshold of 10 µg/g. Among 64 assessments of several different FITs for advanced adenoma, sensitivity ranged from 40% (CI, 33 to 47%) for a threshold of 10 µg/g to 25% (CI, 20 to 31%) for a threshold of 20 µg/g, while specificity ranged from 90% (CI, 87 to 93%) for a threshold of 10 µg/g to 95% (CI, 94 to 96%) for a threshold of ≥ 20 µg/g feces (Figure 3). Positive likelihood ratios ranged from 3.39 at a threshold of < 10 µg/g to 5.58 at a threshold of > 20 µg/g, while the range for negative likelihood ratios was 0.67 to 0.79. Summary receiver operating characteristics curves and 95% confidence regions of the two parameters are displayed in Appendix Figure 2. For thresholds of 10 µg/g, >10 to < 20 µg/g and 20 µg/g, respective areas under the curve were 0.94 for CRC at all three thresholds and were 0.73, 0.62, and 0.69 for advanced adenoma.

Subgroup analyses based on FIT brand and threshold

Subgroup analyses are shown in Appendix Figures 3-6, which are displayed separately for three individual brand FITs for both CRC and advanced adenoma, and in combination for the remaining FITs. Results are most robust for OC Sensor because of the number of assessment made at each of 4 thresholds and for OC Light because of 5 assessments made at its single threshold of 10 µg/g feces. CRC sensitivity of OC Sensor ranged from 73% (CI, 48 to 89%) for a threshold > 20 µg/g to 86% (CI, 75 to 93%) for a threshold of ≤ 10 µg/g, while specificity ranged from 95% (CI, 94 to 96%) to 90% (CI, 86 to 93%) for the same respective thresholds (Appendix Figures 3a). Positive likelihood ratios ranged from 8.45 to 14.71, while negative likelihood ratios ranged from 0.15 to 0.28. Advanced adenoma sensitivity ranged from 33% at the lowest threshold to 23% at the highest threshold, with corresponding specificities of 91% and 95% (Appendix Figure 3b). Positive and negative likelihood ratios had a narrower range for

advanced adenoma: 3.66 to 4.46 and 0.73 to 0.81, respectively. For OC Light's single threshold of 10 µg/g, CRC sensitivity and specificity were 90% and 91%, respectively; for advanced adenoma, respective values were 43% (CI, 24 to 66%) and 91% (CI, 83 to 95%) (Appendix Figures 4a and 4b). OC Light's positive and negative likelihood ratios for CRC were 10.13 and 0.11, respectively. Aggregate point estimates for OC Hemodia, based on 1 or 2 studies, were less robust (Appendix Figures 5a and 5b). For the remaining FITs, test characteristics by FIT threshold are shown in Appendix Figures 6a and 6b.

For thresholds of 10 µg/g, 11-19 µg/g, and 20 µg/g, Table 2 compares test characteristics for five FIT brands, 3 of which are based on 1 or 2 studies. The most robust data at all three thresholds exists for OC sensor. At a threshold of 10 µg/g, sensitivity for both CRC and advanced adenoma were highest and overlapped among the four brands compared. Specificity at this cutoff for both CRC and advanced adenoma was numerically highest for OC Hemodia, with a 95% CI that overlapped only with OC Light. At the threshold between 10 µg/g and 20 µg/g, there were large numerical differences in CRC sensitivity (although with overlapping CIs) among OC Sensor, OC Hemodia, and FOB Gold, while CRC specificity was nearly identical for OC Sensor and FOB Gold (and both higher than for OC Hemodia). Differences among the three brands were smaller for advanced adenoma sensitivity. At a threshold of 20 µg/g, CRC sensitivity was numerically higher for FOB Gold, although CIs overlapped with OC Sensor, while advanced adenoma sensitivity was lower for OC Hemodia. Advanced adenoma specificity at this threshold was nearly identical among the four FITs. This analysis was limited by a small number of studies for some brands and by a comparison that is based on different participants which limits inference about relative performance.

DISCUSSION

This systematic review and meta-analysis quantifies and compares FIT test characteristics for both CRC and advanced adenoma at 5 different thresholds, and for 3 of them, compares test characteristics among different brands of FITs among 31 cross-sectional studies in which screening colonoscopy was the reference standard. We found that positivity threshold has a greater effect on sensitivity and negative likelihood ratios than on specificity and positive likelihood ratios. At a threshold of 10 μg/g, sensitivity for CRC is as high as 91%, specificity is 90%, with positive and negative likelihood ratios of 9.19 and 0.10, respectively, magnitudes that are considered to have clinically important effects on diagnostic threshold (59). Sensitivities for advanced adenoma are much lower, ranging from 25-40%, with more modest likelihood ratios. Based on the number of studies and either comparability or numerical superiority to other FITs, it appears that OC-Sensor (a quantitative FIT) and OC-Light (a qualitative FIT) may be the preferred FITs for hospital- and clinic-based testing, respectively, for large- and small-scale use.

Studies varied by sample size, country, population setting, and age range, as well as FIT used, both the brand of test itself and its threshold, with several studies examining multiple thresholds. All study populations were composed of asymptomatic and largely average risk persons who elected to undergo screening colonoscopy as part of a health promotion / disease prevention program on a local, regional, or national level. Prevalence of both CRC and advanced adenoma varied among studies. This variation was likely related to the age of study participants and perhaps to the country or geographic region where the study was conducted. The exact extent to which study populations and disease prevalence affected study findings is difficult to determine, since FIT brand and threshold values vary as well. At any single threshold, 95% confidence limits overlap for nearly all studies for both CRC and advanced adenoma, suggesting that positivity thresholds, rather than disease prevalence, were associated with individual study test characteristics.

Our quantitative results do not include a single, overall summary estimate of sensitivity and specificity from all studies and all thresholds, as this estimate would have limited clinical utility; while it might represent the “best” single estimate across all studies with their varying thresholds, it would not necessarily apply to any single FIT at any single threshold. Further, this analysis cannot provide optimal test thresholds for CRC screening, as “optimal” requires consideration of other factors, among which are colonoscopy resources available to investigate positive FIT results and the closely-related false-positive rate. For CRC alone, the optimal threshold for a positive test might be between 10 µg/g and 20 µg/g (with a false positive rate of 7%) or ≥ 20 µg/g (with a false positive rate of 5%), the former threshold increasing colonoscopy resources for false-positive results alone by 40%. Both categories have positive likelihood ratios of > 10 and respective negative likelihood ratios of 0.20 and 0.30. Considering both CRC and advanced adenoma, either of the same two categories of test thresholds would appear to be optimal, but require a consideration of the tradeoffs between them.

Our findings are consistent with two prior systematic reviews that informed our study methods (17, 18), one of which quantified performance characteristics for CRC only and included studies with the less accurate and potentially biased surrogate reference standard of two-year follow-up without a CRC diagnosis (17). The other systematic review quantified performance characteristics for both CRC and advanced adenoma, but only included studies of high-risk persons (18). In addition, our searches identified a recent systematic review by Gies and colleagues that assessed seven FIT brands across 22 studies (60). Although less comprehensive than ours and despite inclusion of at least one study that contained persons with previous neoplasia (61), Gies and colleagues found areas under receiver operating characteristics curves similar to our findings, and determined that the large degree of heterogeneity reflected variations in test thresholds (60). Our findings are also consistent with a

recent analysis by Selby and colleagues that considered repeated (i.e., programmatic) testing with FIT, quantifying CRC detection at various test thresholds (62). These investigators found that programmatic sensitivity was greater at lower thresholds, but resulted in a higher number of positive test results per cancer detected.

This study has several limitations that require comment. While we were able to assess several of the QUADAS-2 study quality criteria, a lack of detail in study methods precluded knowing the following for several studies: whether a consecutive or random sample of persons participated; whether FIT results and colonoscopic findings were interpreted independently of one another; and the interval between FIT and colonoscopy. We did not include non-English studies, which could result in language bias, nor did we assess for publication bias. From a clinical perspective, these summary-level performance characteristics apply to one-time testing rather serial testing; therefore, the results do not apply to the serial testing that is recommended in clinical practice. Further, we were unable to determine FIT performance characteristics for proximal and distal lesions separately. Zorzi and colleagues showed lower programmatic FIT-based screening for advanced neoplasia in the proximal colon, highlighting the need to understand the degree of differential FIT performance (63). Statistical heterogeneity was moderate or high for all analyses of specificity and for all analyses for sensitivity except for the 10 μg/g threshold for CRC, for which heterogeneity was low. Given that the subgroup summary estimates show performance characteristics that vary as expected based on positivity threshold, the generally high degree of statistical heterogeneity of this systematic review is more likely due to the large sample sizes of the individual studies than to clinically important variation in study populations, particularly since the outcomes of CRC and advanced adenoma were common to all studies.

This systematic review suggests directions for subsequent research on FIT, the most relevant of which is the need for a head-to-head comparison of different FITs examined at various

thresholds for both CRC and advanced adenoma, and subgrouped by proximal and distal locations of these lesions. While such a study would be challenging logistically and would require a very large sample size, it may be feasible within the framework of a regional or national screening colonoscopy program. The study would ensure use of the same stool sample and standardized pre-analytical conditions. Other studies for consideration include an analysis of programmatic performance characteristics based on the number of rounds of FIT and for prioritizing colonoscopy resources, a quantitative analysis of the yield of FIT when combined with risk factors for advanced colorectal neoplasia.

In conclusion, this systematic review provides new information about the test characteristics of FIT for both CRC and advanced adenoma as a function of test threshold. The findings suggest that FIT may be highly sensitive for CRC in a single application, although at the expense of a high false-positive rate. At high specificity, FIT is moderately sensitive for CRC. While FIT is much less sensitive for advanced adenoma, the natural history of this lesion suggests annual transition rates to CRC in the range of 3-6% (64), implying opportunity to detect this lesion with programmatic screening. Health care systems need to consider both quantity and quality of data for a specific FIT, comparability of its population to the study populations for that particular FIT, and the clinical and economic effects of different test thresholds on colonoscopy and systems resources, as consideration of these factors is required to optimize FIT for the early detection and prevention of CRC.

ACKNOWLEDGEMENTS

None.

FUNDING

Department of Medicine and the Melvin & Bren Simon Cancer Center at the Indiana University

School of Medicine and by the Regenstrief Institute, Inc.

AUTHORS ADDRESSES

Thomas F. Imperiale, MD
1101 W. 10th St.
Regenstrief Institute HSR
Indianapolis, IN 46202
timperia@iu.edu

Rachel N. Gruber, MS
1101 W. 10th St.
Regenstrief Institute HSR
Indianapolis, IN 46202
gruberra@regenstrief.org

Timothy E. Stump, MA
410 W. 10th Street
HITS 3000
Indianapolis, IN 46202
tstump@iu.edu

Thomas Emmett, MD
975 West Walnut St.
IB 100
Indianapolis, IN 46202-5121
temmett@iu.edu

Patrick O. Monahan, PhD
410 W. 10th Street
HITS 3000
Indianapolis, IN 46202
pmonahan@iu.edu

REFERENCES

1. Peery AF, Dellon ES, Lund J, et al. Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterol. 2012;143(5):1179-1187 e1173.
2. Lin JS, Piper MA, Perdue LA, et al. Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2016;315(23):2576-2594.
3. Sharaf RN, Ladabaum U**.** Comparative effectiveness and cost-effectiveness of screening colonoscopy vs. sigmoidoscopy and alternative strategies. Am J Gastroenterol. 2013;108(1):120-32.
4. Heitman SJ, Hilsden RJ, Au F, Dowden S, Manns BJ**.** Colorectal cancer screening for average-risk North Americans: an economic evaluation. PLoS Med. 2010;7(11):e1000370.
5. Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin.* 2018;68(4):250-281.
6. Maxwell AE, Hannon PA, Escoffery C, et al. Promotion and provision of colorectal cancer screening: a comparison of colorectal cancer control program grantees and nongrantees, 2011-2012. Prev Chronic Dis. 2014;11:E170.
7. Navarro M, Nicolas A, Ferrandez A, Lanas A. Colorectal cancer population screening programs worldwide in 2016: An update. World J Gastroenterol. 2017;23(20):3632-3642.
8. Altobelli E, D'Aloisio F, Angeletti PM**.** Colorectal cancer screening in countries of European Council outside of the EU-28. World J Gastroenterol. 2016;22(20):4946-57.
9. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-1477.
10. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med. 1993;328(19):1365-1371.
11. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med. 2013;369(12):1106-1114.
12. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O**.** Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348(9040):1467-71.
13. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. N Engl J Med. 2000;343(22):1603-1607.
14. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut. 2013;62(3):409-415.
15. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology. 2008;135(1):82-90.

16. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59(1):62-68.
17. 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.
18. 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(8):1110-8.
19. Institute of Medicine Committee on Standards for Systematic Reviews of Comparative Effectiveness R. In: Eden J, Levit L, Berg A, Morton S, eds. Finding What Works in Health Care: Standards for Systematic Reviews. Washington (DC): National Academies Press (US) Copyright 2011 by the National Academy of Sciences. All rights reserved.; 2011.
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009 Aug 18;151(4):W65-94.
21. 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(8):529-536.
22. Higgins JP, Thompson SG, Deeks JJ, Altman DG**.** Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
23. Chu H, Cole SR**.** Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol. 2006;59(12):1331-2; author reply 2-3.
24. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A**.** Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol. 2006;6(1):31.
25. Partlett C, Takwoingi Y**.** Meta-analysis of test accuracy studies in R: a summary of user-written programs and step-by-step guide to using glmer. Version 1.0. August 2016. 2018.
26. Bates D, Mächler M, Bolker B, Walker S**.** Fitting linear mixed-effects models using lme4. J Statistical Software. 2015;67(1):1-48.
27. Team RC**.** R: A language and environment for statistical computing. 2013.
28. Nakama H, Zhang B, Zhang X**.** Evaluation of the optimum cut-off point in immunochemical occult blood testing in screening for colorectal cancer. Eur J Cancer. 2001;37(3):398-401.
29. Cheng TI, Wong JM, Hong CF, et al. Colorectal cancer screening in asymptomaic adults: comparison of colonoscopy, sigmoidoscopy and fecal occult blood tests. J Formos Med Assoc. 2002;101(10):685-90.
30. Sohn DK, Jeong SY, Choi HS, et al. Single immunochemical fecal occult blood test for detection of colorectal neoplasia. Cancer Res Treat. 2005;37(1):20-3.
31. Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, Shiratori Y**.** A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. Gastroenterol. 2005;129(2):422-8.

32. Nakazato M, Yamano HO, Matsushita HO, et al. Immunologic fecal occult blood test for colorectal cancer screening. Japan Med Assoc J. 2006;49(5-6):203-7.
33. Graser A, Stieber P, Nagel D, et al. Comparison of CT colonography, colonoscopy, sigmoidoscopy and faecal occult blood tests for the detection of advanced adenoma in an average risk population. Gut. 2009;58(2):241-8.
34. Brenner H, Haug U, Hundt S**.** Inter-test agreement and quantitative cross-validation of immunochromatographical fecal occult blood tests. Int J Cancer. 2010;127(7):1643-9.
35. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol. 2010;105(9):2017-25.
36. Parra-Blanco A, Gimeno-Garcia AZ, Quintero E, et al. Diagnostic accuracy of immunochemical versus guaiac faecal occult blood tests for colorectal cancer screening. J Gastroenterol. 2010;45(7):703-12.
37. Chiang TH, Lee YC, Tu CH, Chiu HM, Wu MS**.** Performance of the immunochemical fecal occult blood test in predicting lesions in the lower gastrointestinal tract. CMAJ. 2011;183(13):1474-81.
38. Haug U, Kuntz KM, Knudsen AB, Hundt S, Brenner H**.** Sensitivity of immunochemical faecal occult blood testing for detecting left- vs right-sided colorectal neoplasia. Br J Cancer. 2011;104(11):1779-85.
39. Khalid-de Bakker CA, Jonkers DM, Sanduleanu S, et al. Test performance of immunologic fecal occult blood testing and sigmoidoscopy compared with primary colonoscopy screening for colorectal advanced adenomas. Cancer Prev Res (Phila). 2011;4(10):1563-71.
40. Omata F, Shintani A, Isozaki M, Masuda K, Fujita Y, Fukui T**.** Diagnostic performance of quantitative fecal immunochemical test and multivariate prediction model for colorectal neoplasms in asymptomatic individuals. Eur J Gastroenterol Hepatol. 2011;23(11):1036-41.
41. de Wijkerslooth TR, Stoop EM, Bossuyt PM, et al. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia. Am J Gastroenterol. 2012;107(10):1570-8.
42. Brenner H, Tao S**.** Superior diagnostic performance of faecal immunochemical tests for haemoglobin in a head-to-head comparison with guaiac based faecal occult blood test among 2235 participants of screening colonoscopy. Eur J Cancer. 2013;49(14):3049-54.
43. Chiu HM, Lee YC, Tu CH, et al. Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clin Gastroenterol Hepatol. 2013;11(7):832-8 e1-2.
44. Ng SC, Ching JY, Chan V, et al. Diagnostic accuracy of faecal immunochemical test for screening individuals with a family history of colorectal cancer. Aliment Pharmacol Ther. 2013;38(7):835-41.
45. Chen YY, Chen TH, Su MY, et al. Accuracy of immunochemical fecal occult blood test for detecting colorectal neoplasms in individuals undergoing health check-ups. Advances in Digestive Medicine. 2014;1(3):74-9.

46. Cubiella J, Castro I, Hernandez V, et al. Diagnostic accuracy of fecal immunochemical test in average- and familial-risk colorectal cancer screening. United European Gastroenterol J. 2014;2(6):522-9.
47. Hernandez V, Cubiella J, Gonzalez-Mao MC, et al. Fecal immunochemical test accuracy in average-risk colorectal cancer screening. World J Gastroenterol. 2014;20(4):1038-47.
48. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med. 2014;370(14):1287-97.
49. Stegeman I, de Wijkerslooth TR, Stoop EM, et al. Combining risk factors with faecal immunochemical test outcome for selecting CRC screenees for colonoscopy. Gut. 2014;63(3):466-71.
50. Aniwan S, Rerknimitr R, Kongkam P, et al. A combination of clinical risk stratification and fecal immunochemical test results to prioritize colonoscopy screening in asymptomatic participants. Gastrointest Endosc. 2015;81(3):719-27.
51. Chang LC, Shun CT, Hsu WF, et al. Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. Clin Gastroenterol Hepatol. 2016;15(6):872-9 e1.
52. Chiu HM, Ching JY, Wu KC, et al. A Risk-Scoring System Combined With a Fecal Immunochemical Test Is Effective in Screening High-Risk Subjects for Early Colonoscopy to Detect Advanced Colorectal Neoplasms. Gastroenterology. 2016;150(3):617-25 e3.
53. Siripongpreeda B, Mahidol C, Dusitanond N, et al. High prevalence of advanced colorectal neoplasia in the Thai population: a prospective screening colonoscopy of 1,404 cases. BMC Gastroenterol. 2016;16:101.
54. Aniwan S, Ratanachu-Ek T, Pongprasobchai S, et al. Impact of Fecal Hb Levels on Advanced Neoplasia Detection and the Diagnostic Miss Rate For Colorectal Cancer Screening in High-Risk vs. Average-Risk Subjects: a Multi-Center Study. Clin Transl Gastroenterol. 2017;8(8):e113.
55. Brenner H, Niedermaier T, Chen H**.** Strong subsite-specific variation in detecting advanced adenomas by fecal immunochemical testing for hemoglobin. Int J Cancer. 2017;140(9):2015-22.
56. Kim NH, Park JH, Park DI, Sohn CI, Choi K, Jung YS**.** The fecal immunochemical test has high accuracy for detecting advanced colorectal neoplasia before age 50. Dig Liver Dis. 2017;49(5):557-61.
57. Shapiro JA, Bobo JK, Church TR, et al. A Comparison of Fecal Immunochemical and High-Sensitivity Guaiac Tests for Colorectal Cancer Screening. Am J Gastroenterol. 2017;112(11):1728-35.
58. Brenner H, Qian J, Werner S**.** Variation of diagnostic performance of fecal immunochemical testing for hemoglobin by sex and age: results from a large screening cohort. Clin Epidemiol. 2018;10:381-9.
59. Grimes DA, Schulz KF**.** Refining clinical diagnosis with likelihood ratios. Lancet. 2005;365(9469):1500-5.
60. Gies A, Bhardwaj M, Stock C, Schrotz-King P, Brenner H**.** Quantitative fecal immunochemical tests for colorectal cancer screening. Int J Cancer. 2018;143(2):234-44.
61. Redwood DG, Asay ED, Blake ID, et al. Stool DNA Testing for Screening Detection of Colorectal Neoplasia in Alaska Native People. Mayo Clin Proc. 2016;91(1):61-70.

62. Selby K, Jensen CD, Lee JK, et al. Influence of Varying Quantitative Fecal Immunochemical Test Positivity Thresholds on Colorectal Cancer Detection: A Community-Based Cohort Study. Ann Intern Med. 2018;169(7):439-47.
63. Zorzi M, Hassan C, Capodaglio G, 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.
64. Brenner H, Hoffmeister M, Stegmaier C, Brenner G, Altenhofen L, Haug U**.** Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut. 2007;56(11):1585-9.

FIGURE LEGEND

Figure 1. Flow diagram of evidence search and study selection

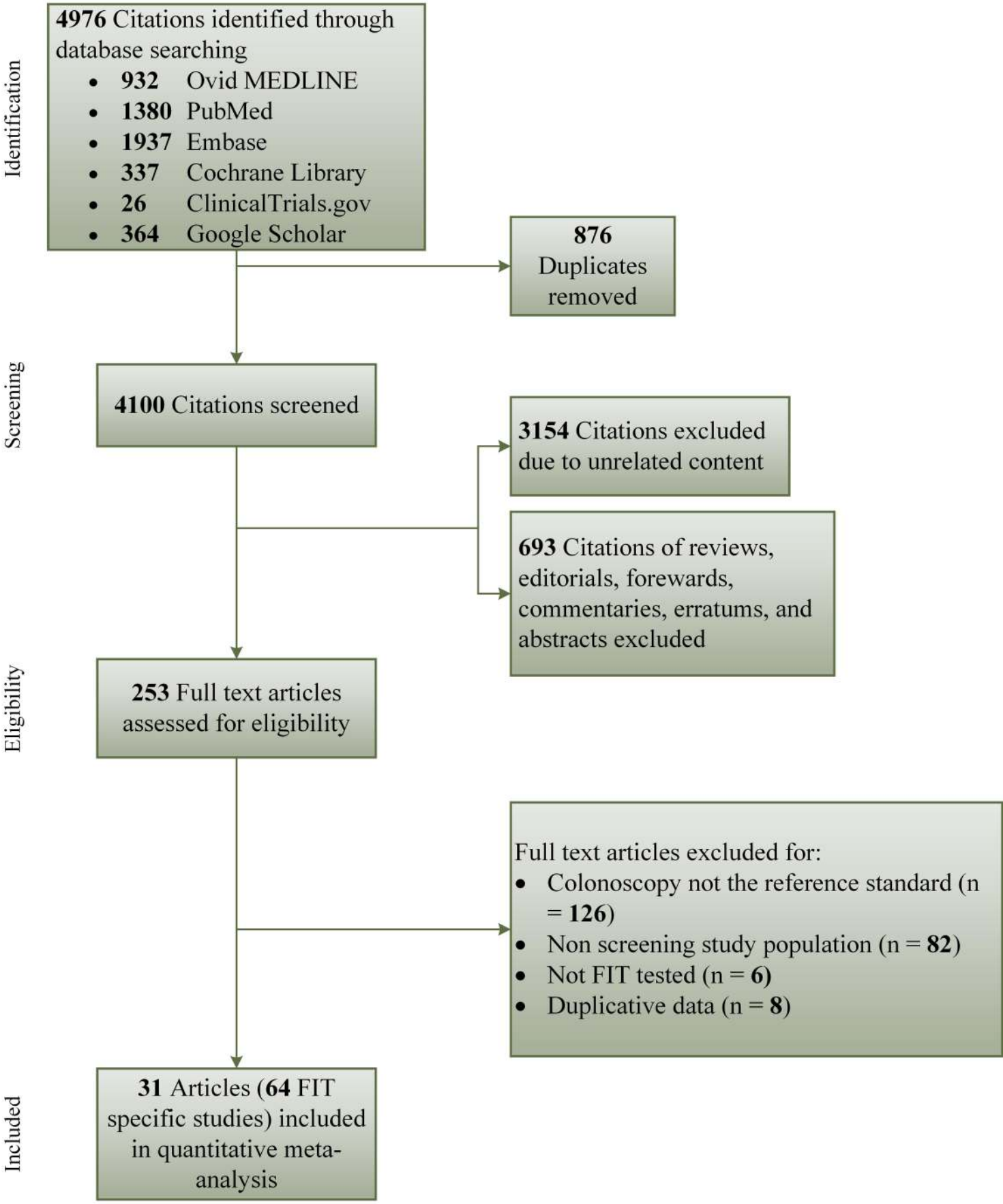
Figure 2. Summary-level Test Characteristics by FIT threshold for Colorectal Cancer

Figure 3. Summary-level Test Characteristics by FIT threshold for Advanced Adenoma

Table 1. Comparison of Brand-specific Test Characteristics at Various Thresholds

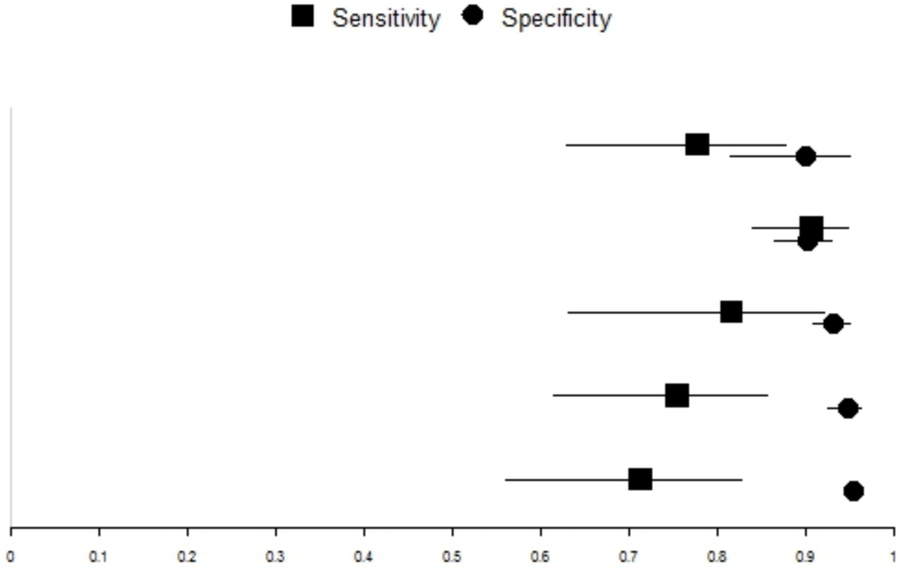
Test	Study N /Patients with CRC	Study N /Patients with AA	CRC Sensitivity (CI)	CRC Specificity (CI)	AA Sensitivity (CI)	AA Specificity (CI)	CRC Positive Likelihood Ratio (95% CI)	CRC Negative Likelihood Ratio (95% CI)	AA Positive Likelihood Ratio (95% CI)	AA Negative Likelihood Ratio (95% CI)
Threshold = 10										
OC Sensor	6 / 56	7 / 898	0.86 (0.75, 0.93)	0.90 (0.86, 0.93)	0.33 (0.30, 0.39)	0.91 (0.86, 0.92)	9.94 (7.07, 12.99)	0.13 (0.06, 0.27)	3.72 (2.87, 4.80)	0.73 (0.69, 0.78)
OC Light	5 / 99	5 / 1027	0.90 (0.72, 0.97)	0.91 (0.83, 0.95)	0.43 (0.24, 0.66)	0.91 (0.83, 0.95)	10.13 (4.34, 21.01)	0.11 (0.03, 0.34)	4.81 (1.41, 14.11)	0.62 (0.36, 0.92)
OC Hemodia	1 / 27	1 / 56	0.89 (0.71, 0.98)	0.94 (0.93, 0.95)	0.59 (0.45, 0.72)	0.94 (0.93, 0.95)	14.62 (10.34, 18.17)	0.12 (0.02, 0.31)	9.69 (6.57, 13.38)	0.44 (0.30, 0.59)
FOB Gold	1 / 25	1 / 286	0.96 (0.80, 1.00)	0.88 (0.87, 0.89)	0.49 (0.43, 0.55)	0.88 (0.87, 0.89)	8.00 (6.02, 9.22)	0.05 (0.00, 0.23)	4.05 (3.22, 5.03)	0.58 (0.51, 0.66)
Threshold >10 <20										
OC Sensor	4 /34	6 / 702	0.81 (0.55, 0.94)	0.93 (0.91, 0.93)	0.29 (0.25, 0.34)	0.93 (0.92, 0.94)	10.93 (6.46, 14.45)	0.20 (0.06, 0.49)	4.41 (3.16, 6.09)	0.76 (0.70, 0.81)
OC Hemodia	1 / 19	1 / 53	0.53 (0.29, 0.76)	0.87 (0.86, 0.89)	0.25 (0.14, 0.38)	0.87 (0.86, 0.89)	4.17 (2.08, 6.59)	0.54 (0.28, 0.83)	1.94 (0.99, 3.34)	0.86 (0.70, 1.00)
FOB Gold	1 / 29	1 / 354	0.97 (0.82, 1.00)	0.94 (0.93, 0.95)	0.37 (0.32, 0.43)	0.97 (0.96, 0.97)	16.45 (12.17, 19.72)	0.04 (0.00, 0.19)	6.35 (4.78, 8.40)	0.67 (0.61, 0.73)
Threshold = 20										
OC Sensor	11 / 163	12 / 2286	0.77 (0.66, 0.85)	0.94 (0.91, 0.96)	0.26 (0.20, 0.32)	0.95 (0.92, 0.96)	13.88 (7.68, 24.03)	0.24 (0.15, 0.37)	4.70 (2.50, 8.74)	0.79 (0.71, 0.87)
OC Hemodia	1 / 12	1 / 67	0.25 (0.06, 0.57)	0.96 (0.96, 0.97)	0.06 (0.02, 0.15)	0.96 (0.96, 0.97)	6.98 (1.30, 19.02)	0.78 (0.44, 0.99)	1.67 (0.39, 4.85)	0.98 (0.88, 1.03)
FOB Gold	1 / 25	1 / 286	0.92 (0.74, 1.00)	0.95 (0.94, 0.96)	0.34 (0.28, 0.40)	0.95 (0.94, 0.96)	17.67 (12.17, 22.37)	0.08 (0.01, 0.28)	6.51 (4.68, 8.97)	0.70 (0.63, 0.76)
Magstream 1000/Hem SP	1 / 79	1 / 648	0.66 (0.54, 0.76)	0.95 (0.95, 0.95)	0.23 (0.19, 0.26)	0.95 (0.95, 0.95)	13.42 (10.43, 16.49)	0.36 (0.25, 0.48)	4.56 (3.69, 5.58)	0.82 (0.78, 0.85)

AA: Advanced Adenoma; CI: 95% Confidence Interval; CRC: colorectal cancer
All results generated using a bivariate model


Identification
Screening
Eligibility
Included
4976 Citations identified through database searching
932 Ovid MEDLINE
1380 PubMed
1937 Embase
337 Cochrane Library
26 ClinicalTrials.gov
364 Google Scholar
876 Duplicates removed
4100 Citations screened
3154 Citations excluded due to unrelated content
693 Citations of reviews, editorials, forewards, commentaries, erratums, and abstracts excluded
253 Full text articles assessed for eligibility
Full text articles excluded for:
Colonoscopy not the reference standard (n = 126)
Non screening study population (n = 82)
Not FIT tested (n = 6)
Duplicative data (n = 8)
31 Articles (64 FIT specific studies) included in quantitative meta-analysis

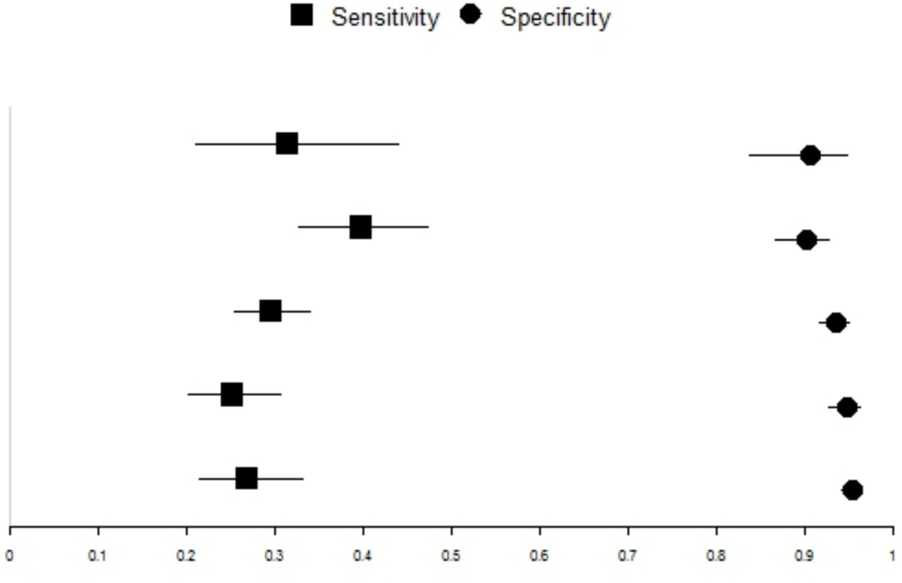
Colorectal Cancer

FIT Threshold µg/g	Study N	Patient N	Patients with CRC	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<10	10	8364	65	0.78 (0.63, 0.88), 0.90 (0.81, 0.95)	7.79 (3.38, 17.19), 0.25 (0.13, 0.46)
=10	16	50892	225	0.91 (0.84, 0.95), 0.90 (0.86, 0.93)	9.19 (6.17, 13.40), 0.10 (0.06, 0.19)
>10 <20	7	12727	95	0.82 (0.63, 0.92), 0.93 (0.91, 0.95)	11.88 (6.80, 18.24), 0.20 (0.08, 0.41)
=20	14	56638	279	0.75 (0.61, 0.86), 0.95 (0.92, 0.96)	14.19 (8.16, 22.95), 0.26 (0.15, 0.42)
>20	12	17341	117	0.71 (0.56, 0.83), 0.95 (0.94, 0.96)	15.49 (9.82, 22.39), 0.30 (0.18, 0.47)



Advanced Adenoma

FIT Threshold µg/g	Study N	Patient N	Patients with AA	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<10	11	8937	769	0.31 (0.21, 0.44), 0.91 (0.84, 0.95)	3.39 (1.31, 8.52), 0.76 (0.59, 0.94)
=10	17	57001	2496	0.40 (0.33, 0.47), 0.90 (0.87, 0.93)	4.05 (2.46, 6.60), 0.67 (0.57, 0.78)
>10 <20	9	19165	1324	0.30 (0.25, 0.34), 0.94 (0.92, 0.95)	4.57 (3.05, 6.82), 0.75 (0.69, 0.81)
=20	15	62747	3287	0.25 (0.20, 0.31), 0.95 (0.93, 0.96)	4.80 (2.81, 8.15), 0.79 (0.72, 0.86)
>20	12	17341	966	0.27 (0.21, 0.33), 0.95 (0.94, 0.96)	5.86 (3.77, 8.97), 0.77 (0.69, 0.83)



TECHNICAL APPENDIX

The approach of Chu and Cole (2006) was used to calculate summary-level estimates of sensitivity and specificity across studies. This approach produces unbiased estimates of sensitivity, specificity and their correlation and does not utilize the ad hoc continuity correction of zero marginal counts of other methods (Reitsma et al, 2005). To automate the process, an R function was created using commands from a tutorial written by Partlett and Takwoingi (2016), obtained from the Cochrane methods website (https://methods.cochrane.org/sdt/software-meta-analysis-dta-studies).

Below is the R function used to obtain estimates for the bivariate generalized mixed models:

```
runmod <- function(dsn) {

   ### Create temporary dataset based upon dsn argument ###
   ### note: dsn must contain TP, FN, FP and TN variables ###
   temp <- as.data.frame(dsn)

   ### Set up the data ###
   ### Generate 5 new variables of type long. We need these before we can reshape the data.
   ### These variables will be included in the glmer function call.
   # n1 is number diseased
   # n0 is number without disease
   # true1 is number of true positives
   # true0 is the number of true negatives
   # study is the unique identifier for each study. _n will generate a sequence of numbers.
   temp$n1 <- temp$TP+temp$FN
   temp$n0 <- temp$FP+temp$TN
   temp$true1 <- temp$TP
   temp$true0 <- temp$TN
   temp$study <- 1:length(temp$n1)

   ### Reshape the data from wide to long format ###
   long <- reshape(temp, direction = "long",
                       varying = list( c("n1" , "n0") , c( "true1","true0" )),
                       timevar = "sens",
                       times = c(1,0),
                       v.names = c("n","true"))

   ### Sort data by study to cluster the 2 records per study together ###
   long <- long[order(long$id),]
   long$spec<- 1-long$sens

   ### Run glmer model to obtain sensitivity and specificity estimates ###
   ### note: between study covariance matrix is unstructured
   y1 <- glmer(formula = cbind(true , n - true) ~ 0 + sens + spec + (0+sens + spec|study),
               data = long,
               glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 100000)),
               family = binomial)

   ### More detail can be obtained by using the summary command ###
   s <- summary(y1)

   ### Extract the coefficients from the model ###
   lsens = s$coeff[1,1]
   lspec = s$coeff[2,1]

   se.lsens = s$coeff[1,2]
   se.lspec = s$coeff[2,2]

   ### Create 95% confidence intervals for logit sens and spec ###
   logit_Sens = c(lsens, lsens-qnorm(0.975)*se.lsens, lsens+qnorm(0.975)*se.lsens )
   logit_Spec = c(lspec, lspec-qnorm(0.975)*se.lspec, lspec+qnorm(0.975)*se.lspec )

   ### R has a built in logit and inv.logit function (use qlogis and plogis) ###
```

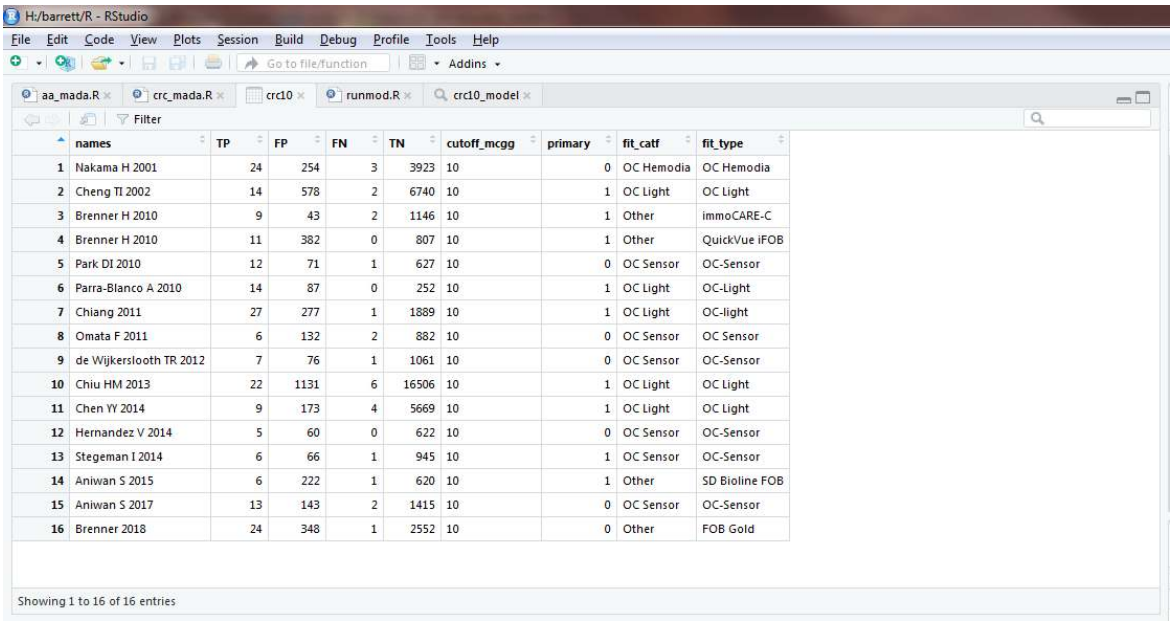
```
   sens <- plogis(logit_Sens)
   spec <- plogis(logit_Spec)

   ### Create data frame containing sens/spec estimates ###
   f <- t(data.frame(logit_Sens, sens, logit_Spec, spec))
   colnames(f) <- c("mean","low","hi")
   rownames(f) <- c("logit_Sens","sens","logit_Spec","spec")

   ### Return a list containing model summary and calculated sens/spec estimates ###
   l <- list(s, f)
   return(l)
}
```

To run this function, you will need to make sure the lme4 library is loaded by running the following statement in your R script: `library(lme4)`

An example call to this function would be:

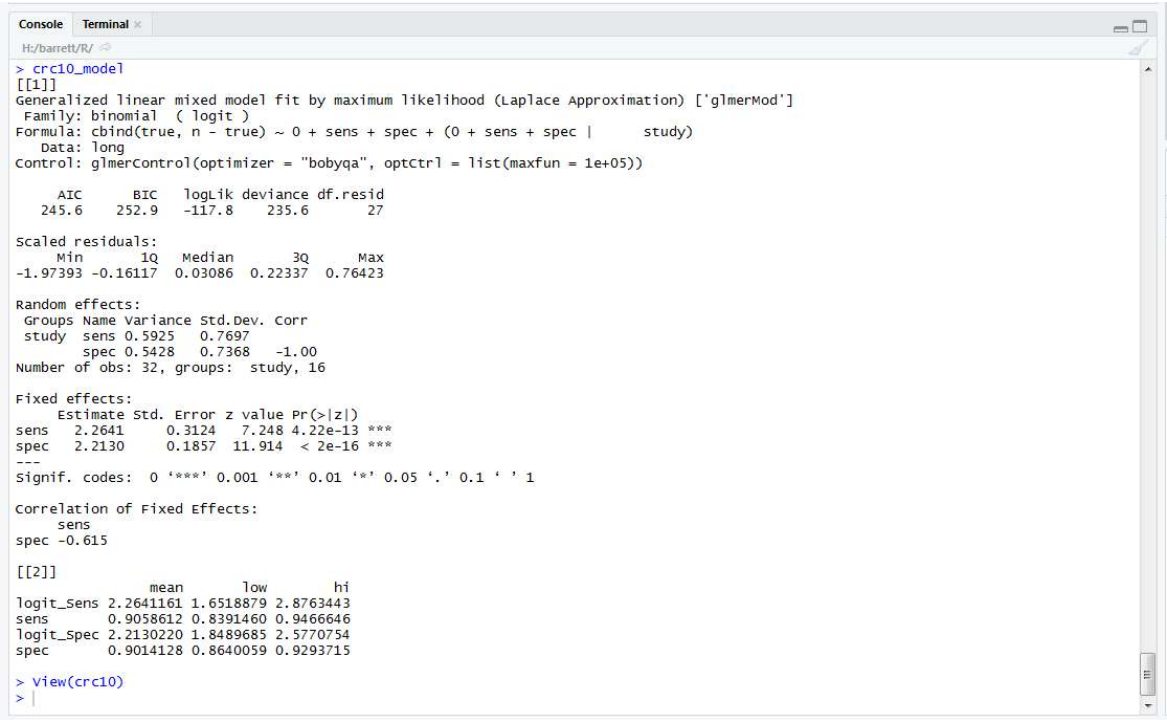
```
crc10_model <- runmod(crc10)
```

where `crc10` is a data frame containing the TP, FP, FN and TN variables. An example would be the following:



	names	TP	FP	FN	TN	cutoff_mcgg	primary	fit_catf	fit_type
1	Nakama H 2001	24	254	3	3923	10	0	OC Hemodia	OC Hemodia
2	Cheng TI 2002	14	578	2	6740	10	1	OC Light	OC Light
3	Brenner H 2010	9	43	2	1146	10	1	Other	immoCARE-C
4	Brenner H 2010	11	382	0	807	10	1	Other	QuickVue iFOB
5	Park DI 2010	12	71	1	627	10	0	OC Sensor	OC-Sensor
6	Parra-Blanco A 2010	14	87	0	252	10	1	OC Light	OC-Light
7	Chiang 2011	27	277	1	1889	10	1	OC Light	OC-light
8	Omata F 2011	6	132	2	882	10	0	OC Sensor	OC Sensor
9	de Wijkerslooth TR 2012	7	76	1	1061	10	0	OC Sensor	OC-Sensor
10	Chiu HM 2013	22	1131	6	16506	10	1	OC Light	OC Light
11	Chen YY 2014	9	173	4	5669	10	1	OC Light	OC Light
12	Hernandez V 2014	5	60	0	622	10	0	OC Sensor	OC-Sensor
13	Stegeman I 2014	6	66	1	945	10	1	OC Sensor	OC-Sensor
14	Aniwan S 2015	6	222	1	620	10	1	Other	SD Bioline FOB
15	Aniwan S 2017	13	143	2	1415	10	0	OC Sensor	OC-Sensor
16	Brenner 2018	24	348	1	2552	10	0	Other	FOB Gold

The returned object from the function is a list containing the model summary as the first element and the sens/spec estimates as the second element. An example of the output returned would be the following:



References:

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of clinical epidemiology.* 2006;59(12):1331-1332; author reply 1332-1333.

Partlett C, Takwoingi Y. Meta-analysis of test accuracy studies in R: a summary of user-written programs and step-by-step guide to using glmer. Version 1.0. August 2016. Available from: http://methods.cochrane.org/sdt/.

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. *Journal of Clinical Epidemiology.* 2005;58(10):982-990.

Appendix Table of Contents

Appendix Table 1: Detailed Search Strategies

MEDLINE (Ovid):

1	exp Feces/
2	exp Occult Blood/
3	(feces or faeces or fecal or faecal or stool or occult blood or fob*).tw.
4	exp Colorectal Neoplasms/
5	(colon or colonic or colorectal or colo rectal).tw.
6	exp Mass Screening/
7	exp "Early Detection of Cancer"/
8	screen*.tw.
9	(or/4-5) and (or/6-8)
10	Immunochemistry/
11	Immunologic Tests/
12	(immunochem* or immuno chem*).tw.
13	fit.ti.
14	(or/1-3,9) and (or/10-13)
15	((feces or faeces or fecal or faecal or stool or occult blood or fob*) adj10 (immunologic* or immunochromatograph* or immuno chromatograph* or immunohistochem* or immuno histochem*)).tw.
16	(ifobt or i fobt or immunofecal* or immuno fecal* or immunofaecal* or immuno faecal*).tw.
17	(hemeselect or heme select or hemoccultsensa or hemoccult sensa or immudia or magstream or bayer detect or flexsure or oc auto or (monohaem not cytochrome*) or oc sensor or hemodia or fobgold or sentifob).tw.
18	or/14-17
19	limit 18 to english language
20	exp Animals/ not exp Humans/
21	19 not 20

PubMed (PubMed.gov):

#1	feces [tiab] OR faeces [tiab] OR fecal [tiab] OR faecal [tiab] OR stool [tiab] OR "occult blood" [tiab] OR fob* [tiab]
#2	(colon [tiab] OR colonic [tiab] OR colorectal [tiab] OR "colo rectal" [tiab]) AND screen* [tiab]
#3	immunochem* [tiab] OR immuno chem* [tiab] OR immunologic* [tiab] OR immunochromatograph* [tiab] OR immuno chromatograph* [tiab] OR immunohistochem* [tiab] OR immuno histochem* [tiab] OR fit [ti]
#4	ifobt [tiab] OR "i fobt" [tiab] OR immunofecal* [tiab] OR immuno fecal* [tiab] OR immunofaecal* [tiab] OR immuno faecal* [tiab]
#5	hemeselect [tiab] OR "heme select" [tiab] OR hemoccultsensa [tiab] OR "hemoccult sensa" [tiab] OR immudia [tiab] OR magstream [tiab] OR "bayer detect" [tiab] OR flexsure [tiab] OR "oc auto" [tiab] OR (monohaem [tiab] NOT cytochrome* [tiab]) OR "oc sensor" [tiab] OR hemodia [tiab] OR fobgold [tiab] OR sentifob [tiab]
#6	Search ((#1 OR #2) AND #3) OR #4 OR #5 NOT medline [sb] Filters: English

EMBASE (Embase.com):

#1	'feces'/exp
#2	'occult blood'/exp

#3 feces:ab,ti OR faeces:ab,ti OR fecal:ab,ti OR faecal:ab,ti OR stool:ab,ti OR 'occult blood':ab,ti OR fob*:ab,ti
#4 'colon cancer'/exp
#5 colon:ab,ti OR colonic:ab,ti OR colorectal:ab,ti OR 'colo rectal':ab,ti
#6 'cancer screening'/exp
#7 'early diagnosis'/exp
#8 screen*:ab,ti
#9 (#4 OR #5) AND (#6 OR #7 OR #8)
#10 'immunochemistry'/de
#11 'immunological procedures'/de
#12 immunochem*:ab,ti OR (immuno NEXT/1 chem*):ab,ti
#13 fit:ti
#14 (#1 OR #2 OR #3 OR #9) AND (#10 OR #11 OR #12 OR #13)
#15 ((feces OR faeces OR fecal OR faecal OR stool OR 'occult blood' OR fob*) NEAR/10 (immunologic* OR immunochromatograph* OR chromatograph* OR immunohistochem* OR histochem*)):ab,ti
#16 ifobt:ab,ti OR 'i fobt':ab,ti OR immunofecal*:ab,ti OR 'immuno fecal':ab,ti OR immunofaecal*:ab,ti OR 'immuno faecal':ab,ti
#17 hemeselect:ab,ti OR 'heme select':ab,ti OR hemoccultsensa:ab,ti OR 'hemoccult sensa':ab,ti OR immudia:ab,ti OR magstream:ab,ti OR 'bayer detect':ab,ti OR flexsure:ab,ti OR 'oc auto':ab,ti OR 'oc sensor':ab,ti OR hemodia:ab,ti OR fobgold:ab,ti OR sentifob:ab,ti
#18 monohaem:ab,ti NOT cytochrome*:ab,ti
#19 #14 OR #15 OR #16 OR #17 OR #18
#20 'animal'/exp NOT 'human'/exp
#21 (#19 NOT #20) AND [english]/lim

Cochrane Library (Wiley):

#1 (feces or faeces or fecal or faecal or stool or "occult blood" or fob*):ti,ab,kw
#2 ((colon or colonic or colorectal or "colo rectal") and screen*):ti,ab,kw
#3 (immunochem* or immuno chem* or immunologic* or immunochromatograph* or immuno chromatograph* or immunohistochem* or immuno histochem*):ti,ab,kw or fit:ti
#4 (ifobt or "i fobt" or immunofecal* or immuno fecal* or immunofaecal* or immuno faecal* or hemeselect or "heme select" or hemoccultsensa or "hemoccult sensa" or immudia or magstream or "bayer detect" or flexsure or "oc auto" or (monohaem not cytochrome*) or "oc sensor" or hemodia or fobgold or sentifob):ti,ab,kw
#5 ((#1 or #2) and #3) or #4

Appendix Table 2. Selected Characteristics of Included Studies

Study, Year (ref)	Study Type	Population Setting	Funding Source	Cohort size, n	FIT brand	Mean age (range)	FIT samples, n	µg/g	Timing of FIT relative to colonoscopy	FIT interpreted independent of colonoscopy	CRC Prevalence	AA Prevalence	CRC Cases, n	AA Cases, n	CRC Sensitivity (95% CI)	CRC Specificity (95% CI)	AA Sensitivity (95% CI)	AA Specificity (95% CI)
Nakama 2001 (28)	Prospective cross sectional	Hospital based CRC screening program	Ministry of Health and Welfare of Japan	4260	OC Hemodia	NA (40 to 70+)	2	10	Prior	Unknown	0.63%	1.31%	27	56	0.89 (0.71, 0.98)	0.94 (0.93, 0.95)	0.59 (0.45, 0.72)	0.94 (0.93, 0.95)
								30							0.82 (0.62, 0.94)	0.96 (0.96, 0.97)	0.54 (0.40, 0.67)	0.96 (0.96, 0.97)
								60							0.56 (0.35, 0.75)	0.97 (0.97, 0.98)	0.21 (0.12, 0.34)	0.97 (0.97, 0.98)
Cheng 2002 (29)	Retrospective cross sectional	Hospital based health screening program	Not reported	7411	OC Light	46.8 (≤20 to ≥81)*	1	10	Prior	Yes	0.22%	1.04%	16	77	0.88 (0.62, 0.98)	0.92 (0.91, 0.93)	0.40 (0.29, 0.52)	0.92 (0.91, 0.93)
Sohn 2004 (30)†	Prospective cross sectional	Single center cancer prevention program	Not reported	3794	OC Hemodia	48.9 (15 to 78)	1	20	Prior	Yes	0.32%	1.77%	12	67	0.25 (0.06, 0.57)	0.96 (0.96, 0.97)	0.06 (0.02, 0.15)	0.96 (0.96, 0.97)
Morikawa 2005 (31)	Retrospective cross sectional	Single center cancer prevention program	Not reported	21805	Magstream 1000/Hem SP	48.2 (21 to 91)	1	20	Prior	Yes	0.36%	2.97%	79	648	0.66 (0.54, 0.76)	0.95 (0.95, 0.95)	0.23 (0.19, 0.26)	0.95 (0.95, 0.95)
Nakazato 2006 (32)	Prospective cross sectional	Hospital based CRC screening program	Not reported	3090	OC Hemodia	53.4 (Not provided)	2	16	Prior	Yes	0.61%	1.72%	19	53	0.53 (0.29, 0.76)	0.87 (0.86, 0.89)	0.25 (0.14, 0.38)	0.87 (0.86, 0.89)
Graser 2009 (33)	Prospective cross sectional	Hospital based CRC screening study	Not reported	284	FOB Gold	60.5 (50 to 81)	2	3	Prior	Unknown	NA	8.10%	NA	23	NA	NA	0.30 (0.13, 0.53)	0.86 (0.81, 0.90)
Brenner 2010 (34)	Prospective cross sectional	Regional colonoscopy screening program	German Research Foundation	1330	PreventID	63 (Not provided)	1	2	Prior	Yes	0.83%	9.77%	11	130	1.00 (0.72, 1.00)	0.81 (0.79, 0.83)	0.49 (0.40, 0.58)	0.81 (0.79, 0.84)
					Bionexia Hb/Hp Complex			5							0.82 (0.48, 0.98)	0.56 (0.54, 0.59)	0.72 (0.63, 0.79)	0.56 (0.54, 0.59)
					Bionexia			8							1.00 (0.72, 1.00)	0.80 (0.77, 0.82)	0.52 (0.43, 0.61)	0.80 (0.77, 0.82)
					FOB advanced			8							0.73 (0.39, 0.94)	0.91 (0.90, 0.93)	0.27 (0.20, 0.35)	0.91 (0.90, 0.93)
					immoCARE-C			10							0.82 (0.48, 0.98)	0.96 (0.95, 0.97)	0.25 (0.18, 0.34)	0.96 (0.95, 0.97)
					QuickVue iFOB			10							1.00 (0.72, 1.00)	0.68 (0.65, 0.70)	0.56 (0.47, 0.65)	0.68 (0.65, 0.71)
Park 2010 (35)	Prospective cross sectional	Screening colonoscopy at 4 tertiary centers	Eiken Chemical & Shinyong Diagnostics‡	770	OC Sensor	59.3 (50 to 75)	3	10	Prior	Yes	1.69%	7.66%	13	59	0.92 (0.64, 1.00)	0.90 (0.87, 0.92)	0.44 (0.31, 0.58)	0.90 (0.87, 0.93)
								15							0.92 (0.64, 1.00)	0.91 (0.90, 0.93)	0.37 (0.25, 0.51)	0.91 (0.90, 0.93)
								20							0.92 (0.64, 1.00)	0.92 (0.90, 0.93)	0.34 (0.22, 0.47)	0.92 (0.90, 0.94)
								25							0.85 (0.55, 0.98)	0.93 (0.91, 0.95)	0.29 (0.18, 0.42)	0.93 (0.91, 0.95)
								30							0.85 (0.55, 0.98)	0.94 (0.92, 0.95)	0.27 (0.16, 0.40)	0.94 (0.92, 0.95)
Parra-Blanco 2010 (36)	Prospective cross sectional	Random sample - population based screening	In part by Government Grants	402	OC Light	NA (50 to 79)	1	10	Prior	Yes	3.48%	12.19%	14	49	1.00 (0.77, 1.00)	0.74 (0.69, 0.79)	0.86 (0.73, 0.94)	0.74 (0.69, 0.79)
Chiang 2011 (37)	Prospective cross sectional	Hospital based advertised screening program	Taipei Institute of Pathology	2222	OC Light	49 (19 – 84)	1	10	Prior	Yes	1.26%	1.26%	28	28	0.96 (0.82, 1.00)	0.87 (0.86, 0.89)	0.46 (0.28, 0.66)	0.87 (0.86, 0.87)
Haug 2011 (38) ††	Prospective cross sectional	Regional colonoscopy screening program	German Research Foundation‡	2325	RIDASCREEN Hemo	NA (Not provided)	1	8	Prior	Yes	0.60%	9.20%	13	215	0.77 (0.46, 0.95)	0.95 (0.94, 0.96)	0.27 (0.21, 0.33)	0.95 (0.94, 0.96)
								15							0.77 (0.46, 0.95)	0.97 (0.96, 0.98)	0.26 (0.20, 0.32)	0.97 (0.96, 0.98)
Khalid-de Bakkar, 2011 (39)	Prospective cross sectional	University hospital based screening program	Not reported	329	OC Sensor	54.6 (50 to 65)	1	9	Prior	Unknown	NA	11.55%	NA	38	NA	NA	0.16 (0.06, 0.31)	0.97 (0.94, 0.99)
								15							NA	NA	0.28 (0.15, 0.46)	0.95 (0.92, 0.97)
Omata 2011 (40)	Retrospective cross sectional	University hospital based prevention clinic	Not reported	1085	OC Sensor	64 (Not provided)	1	5	Prior	Yes	0.74%	5.81%	8	63	0.75 (0.35, 0.97)	0.77 (0.74, 0.80)	0.48 (0.35, 0.61)	0.77 (0.74, 0.80)
								10							0.75 (0.35, 0.97)	0.87 (0.85, 0.89)	0.33 (0.22, 0.46)	0.87 (0.85, 0.89)
								15							0.50 (0.16, 0.84)	0.92 (0.90, 0.94)	0.19 (0.10, 0.31)	0.92 (0.90, 0.94)
								20							0.50 (0.16, 0.84)	0.94 (0.92, 0.95)	0.19 (0.10, 0.31)	0.94 (0.92, 0.95)
								25							0.25 (0.03, 0.65)	0.96 (0.95, 0.97)	0.19 (0.10, 0.31)	0.96 (0.95, 0.97)
								30							0.25 (0.03, 0.65)	0.97 (0.96, 0.98)	0.19 (0.10, 0.31)	0.97 (0.96, 0.98)
de Wijkerslooth 2012 (41)	Prospective cross sectional	Random sample - population based screening	Netherlands Research and Development & Center for Translational Molecular Medicine	1256	OC Sensor	NA (Not provided)	1	10	Prior	Yes	0.64%	8.84%	8	111	0.88 (0.47, 1.00)	0.93 (0.92, 0.95)	0.34 (0.26, 0.44)	0.93 (0.92, 0.95)
								15							0.75 (0.35, 0.97)	0.94 (0.92, 0.95)	0.30 (0.21, 0.39)	0.94 (0.92, 0.95)
								20							0.75 (0.35, 0.97)	0.94 (0.93, 0.95)	0.28 (0.20, 0.37)	0.94 (0.93, 0.95)
Brenner 2013 (42)	Prospective cross sectional	Regional colonoscopy screening program	German Research Foundation & Federal Ministry of	2235	RIDASCREEN Haemo/Haptoglobin	NA (Not provided)	1	2	Prior	Yes	0.67%	9.26%	15	207	0.53 (0.27, 0.79)	0.97 (0.96, 0.98)	0.18 (0.13, 0.24)	0.97 (0.96, 0.98)
					RIDASCREEN Hemo			2							0.60 (0.32, 0.84)	0.97 (0.96, 0.98)	0.21 (0.16, 0.27)	0.97 (0.96, 0.98)

			Education and Research		OC Sensor			20							0.73 (0.45, 0.92)	0.97 (0.97, 0.98)	0.22 (0.17, 0.29)	0.97 (0.97, 0.98)
Chiu 2013 (43)	Prospective cross sectional	University hospital based prevention clinic	Department of Health of Taiwan	18297	OC Light	59.8 (50 to 70+)	1	10	Prior	Yes	0.15%	3.45%	28	632	0.79 (0.59, 0.92)	0.94 (0.93, 0.94)	0.28 (0.25, 0.32)	0.94 (0.93, 0.94)
Ng 2013 (44)	Prospective cross sectional	University hospital based prevention clinic	Hong Kong Jockey Club	4539	FIT Hemosure	57.68 (50 to 70)	1	50	Prior	Yes	0.48%	4.82%	22	219	0.55 (0.32, 0.76)	0.91 (0.90, 0.92)	0.39 (0.32, 0.46)	0.91 (0.90, 0.92)
Chen 2014 (45)	Retrospective cross sectional	Hospital based CRC screening program	Not reported	6096	OC Light	53.65 (40 to 87)	1	10	Prior	Yes	0.21%	3.95%	13	241	0.69 (0.39, 0.91)	0.97 (0.97, 0.98)	0.20 (0.15, 0.25)	0.97 (0.97, 0.98)
Cubiella 2014 (46) ††	Retrospective cross sectional	3 Tertiary care hospitals in Spain	Government funded‡	722	OC Sensor	56.9 (50 to 69)	1	20	Prior	Yes	0.42%	11.63%	3	84	1.00 (0.29, 1.00)	0.96 (0.94, 0.98)	0.27 (0.18, 0.38)	0.96 (0.94, 0.98)
Hernandez 2014 (47)	Prospective cross sectional	3 Tertiary care hospitals in Spain	Government funded‡	779	OC Sensor	57.55 (50 to 69)	2	10	Prior	Yes	0.64%	11.81%	5	92	1.00 (0.48, 1.00)	0.91 (0.89, 0.93)	0.39 (0.29, 0.50)	0.91 (0.89, 0.93)
								15							1.00 (0.48, 1.00)	0.93 (0.90, 0.95)	0.37 (0.27, 0.48)	0.93 (0.90, 0.95)
								20							1.00 (0.48, 1.00)	0.93 (0.91, 0.95)	0.34 (0.24, 0.44)	0.93 (0.91, 0.95)
								23							1.00 (0.48, 1.00)	0.94 (0.91, 0.95)	0.33 (0.23, 0.43)	0.94 (0.91, 0.95)
								30							0.80 (0.84, 1.00)	0.94 (0.92, 0.96)	0.29 (0.20, 0.40)	0.94 (0.92, 0.96)
Imperiale 2014 (48)	Prospective cross sectional	90, private practice and academic sites in US	Exact Sciences	9989	OC-FIT-CHEK	NA (50 to ≥75)	1	20	Prior	Yes	0.65%	7.58%	65	757	0.74 (0.62, 0.84)	0.95 (0.94, 0.95)	0.24 (0.21, 0.27)	0.95 (0.94, 0.95)
Stegeman 2014 (49)	Prospective cross sectional	Random sample, population based	Dutch Ministry of Health	1112	OC Sensor	60.6 (50 to 75)	1	10	Unknown	Unknown	0.63%	8.45%	7	94	0.86 (0.42, 1.00)	0.94 (0.92, 0.95)	0.32 (0.23, 0.42)	0.94 (0.92, 0.95)
Aniwan 2015 (50)	Prospective cross sectional	University hospital-based screening program	University Endowment Fund	948	SD Bioline FOB	60.6 (50 to 75)	1	10	Prior	Yes	0.74%	10.44%	7	99	0.86 (0.42, 1.00)	0.74 (0.71, 0.77)	0.49 (0.38, 0.59)	0.74 (0.71, 0.77)
Chang 2016 (51)	Prospective cross sectional	University hospital-based screening program	Ministry of Health and Welfare of Taiwan	6109	OC Sensor	59 (50 to ≥70)	1	10	Prior	Yes	NA	5.55%	NA	339	NA	NA	0.32 (0.28, 0.38)	0.91 (0.90, 0.92)
								15							NA	NA	0.25 (0.20, 0.29)	0.95 (0.94, 0.95)
								20							NA	NA	0.21 (0.17, 0.26)	0.96 (0.96, 0.97)
Chiu 2016 (52)	Prospective cross sectional	Multinational	Not reported	3958	OC Sensor	57.8 (40 to ≥70)	1	20	Prior	Yes	0.45%	5.26%	18	208	0.94 (0.73, 1.00)	0.68 (0.66, 0.69)	0.61 (0.54, 0.67)	0.68 (0.66, 0.69)
Siripongpreeda 2016 (53)	Prospective cross sectional	Hospital based screening program	Hospital Research Grant	1404	FOB one-step	56.9 (50 to 65)	1	6	Prior	Yes	1.28%	6.62%	18	93	0.56 (0.31, 0.79)	0.96 (0.95, 0.97)	0.07 (0.02, 0.14)	0.96 (0.95, 0.97)
Aniwan 2017 (54)†	Prospective cross sectional	6 University hospital based health promotion programs in Thailand	National Research Council of Thailand, Health Systems Research Institute & International Research Integration Grant	1713	OC Sensor	59.4 (50 to 75)	1	10	Prior	Yes	0.88%	8.17%	15	140	0.87 (0.60, 0.98)	0.91 (0.89, 0.92)	0.26 (0.19, 0.34)	0.91 (0.89, 0.92)
								20							0.80 (0.52, 0.96)	0.95 (0.94, 0.96)	0.18 (0.12, 0.25)	0.95 (0.94, 0.96)
								30							0.80 (0.52, 0.96)	0.96 (0.95, 0.97)	0.11 (0.07, 0.18)	0.96 (0.95, 0.97)
Brenner 2017 (55)	Prospective cross sectional	Regional colonoscopy screening program	German Research Council	3437	FOB Gold	NA (50 to 79)	1	17	Prior	Yes	0.84%	10.30%	29	354	0.97 (0.82, 1.00)	0.94 (0.93, 0.95)	0.37 (0.32, 0.43)	0.94 (0.93, 0.95)
Kim 2017 (56)	Retrospective cross sectional	Health Center Screening Program, 2 centers in Korea	Not reported	4374	OC Sensor	NA (Not provided)	1	20	Prior	Unknown	0.25%	4.00%	11	175	0.63 (0.31, 0.89)	0.97 (0.96, 0.97)	0.19 (0.14, 0.26)	0.97 (0.96, 0.97)
Shapiro 2017 (57)	Prospective cross sectional	2 University hospital-based screening programs	Center for Disease Control & Prevention	947	OC FIT-CHEK	NA (50 to 75)	1	20	Prior	Yes	0.21%	5.39%	2	51	0.00 (0.00, 0.84)	0.98 (0.97, 0.99)	0.16 (0.07, 0.29)	0.98 (0.97, 0.99)
				984	InSure FIT		2				0.20%	5.18%			0.50 (0.01, 0.99)	0.97 (0.96, 0.98)	0.26 (0.14, 0.40)	0.97 (0.95, 0.98)
Brenner 2018 (58)	Prospective cross sectional	Regional colonoscopy screening program	German Research Council & Federal Ministry of Education and Research	3211	FOB Gold	NA (50 to 79)	1	10	Prior	Yes	0.78%	8.91%	25	286	0.96 (0.80, 1.00)	0.88 (0.87, 0.89)	0.48 (0.42, 0.54)	0.88 (0.87, 0.89)
								20							0.92 (0.74, 0.99)	0.95 (0.94, 0.96)	0.34 (0.28, 0.40)	0.94 (0.94, 0.96)
								30							0.88 (0.69, 0.97)	0.97 (0.96, 0.97)	0.28 (0.23, 0.34)	0.97 (0.96, 0.97)

AA: Advanced Adenoma; CI: 95% Confidence Interval; CRC: colorectal cancer; FIT: fecal immunochemical test; NA: Not Applicable; µg/g: microgram of hemoglobin per gram of stool
*6 people in the ≤20 age group, 3 in the ≥81 age group
†Numbers of CRCs and AAs calculated from percentages
‡Sponsor had no role in design, analysis, or in manuscript preparation
†† Received communication from authors with absolute numbers

Appendix Table 3. I^2 Values for Sensitivity and Specificity

FIT Threshold μg/g	Colorectal Cancer	Advanced Adenomas
	I^2 Sensitivity, Specificity	I^2 Sensitivity, Specificity
<10	61.5%, 99.5%	95.6%, 99.4%
=10	21.1%, 99.0%	91.2%, 98.9%
>10 <20	72.2%, 96.9%	66.0%, 96.5%
=20	68.7%, 99.5%	91.7%, 99.5%
>20	66.6%, 95.4%	81.2%, 95.4%

Appendix Table 4. Frequencies of Indeterminate FIT Results and Indeterminate Colonoscopies

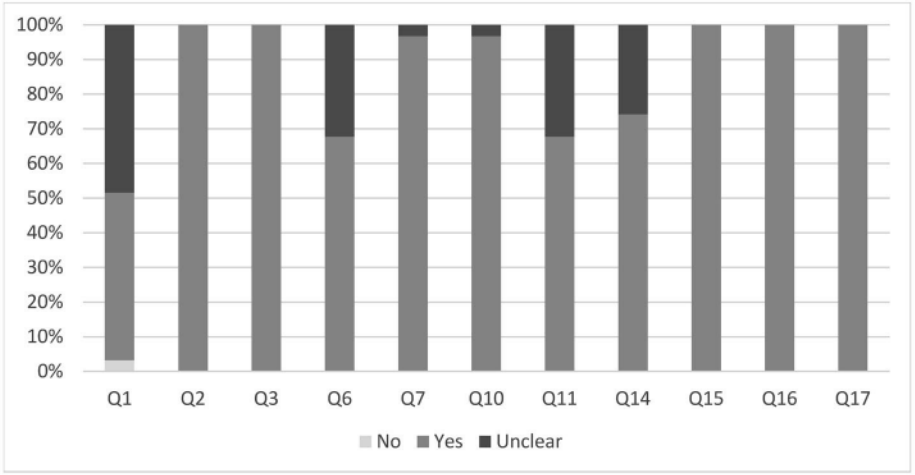
Study, Year (ref)	Indeterminate FIT Results	Indeterminate colonoscopies*
Nakama, 2001 (28)	Not provided	Not provided
Cheng, 2002 (29)	Not provided	Not provided
Sohn, 2004 (30)	Nor provided	Not provided
Morikawa, 2005 (31)	Not provided	1.8%
Nakazato, 2006 (32)	Not provided	1.4%
Graser, 2009 (33)	Not provided	0.64%
Brenner, 2010 (34)	Not provided	Not provided
Park, 2010 (35)	Not provided	0.12%
Parra-Blanco, 2010 (36)	Not provided	4.1%
Chiang, 2011 (37)	Not provided	0.64%
Haug, 2011 (38)	2.0% "without FOBT result"	5.5%
Khalid-de Bakkar, 2011 (39)	Not provided	Not provided†
Omata, 2011 (40)	7.1% "unavailable"	0.93%
de Wijkerslooth, 2012 (41)	Not provided	Not provided
Brenner, 2013 (42)	Not provided	Not provided
Chiu, 2013 (43)	Not provided	1.7%
Ng, 2013 (44)	Not provided	Not provided
Chen, 2014 (45)	Not provided	Not provided‡
Cubiella, 2014 (46)	Not provided	Not provided
Hernandez, 2014 (47)	Not provided	Not provided
Imperiale, 2014 (48)	0.31% insufficient hemoglobin	1.76%
Stegeman, 2014 (49)	Not provided	Not provided
Aniwan, 2015 (50)	Not provided	0.95%
Chang, 2016 (51)	Not provided	1.98%
Chiu, 2016 (52)	Not provided	Not provided
Siripongpreeda, 2016 (53)	Not provided	Not provided
Aniwan, 2017 (54)	Not provided	0.7%
Brenner, 2017 (55)	Not provided	Not provided
Kim, 2017 (56)	Not provided	8.9%
Shapiro, 2017 (57)	Not provided	Not provided
Brenner, 2018 (58)	Not provided	0.9%

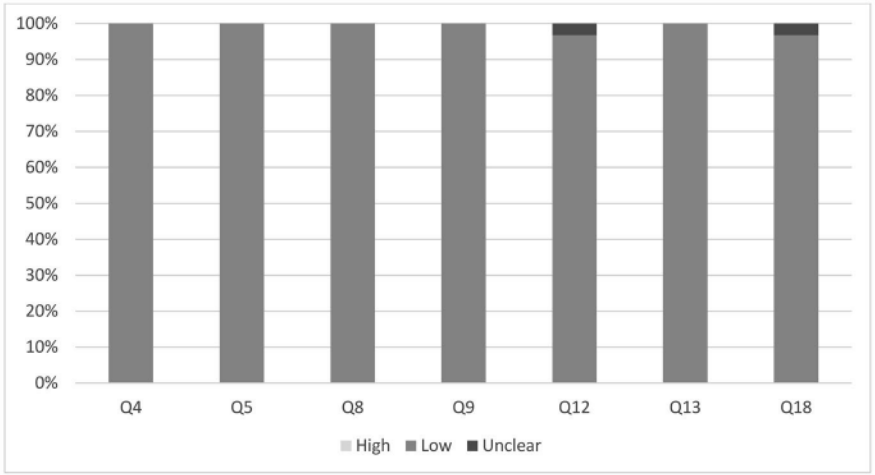
*Due to poor prep quality and/or incomplete colonoscopy

† 118 (26.4%) of 447 participants were excluded because of unavailable FIT results, colonoscopy that was either incomplete or of poor quality preparation, or both reasons.

‡ 8.7% excluded from analysis due to incomplete or missing information

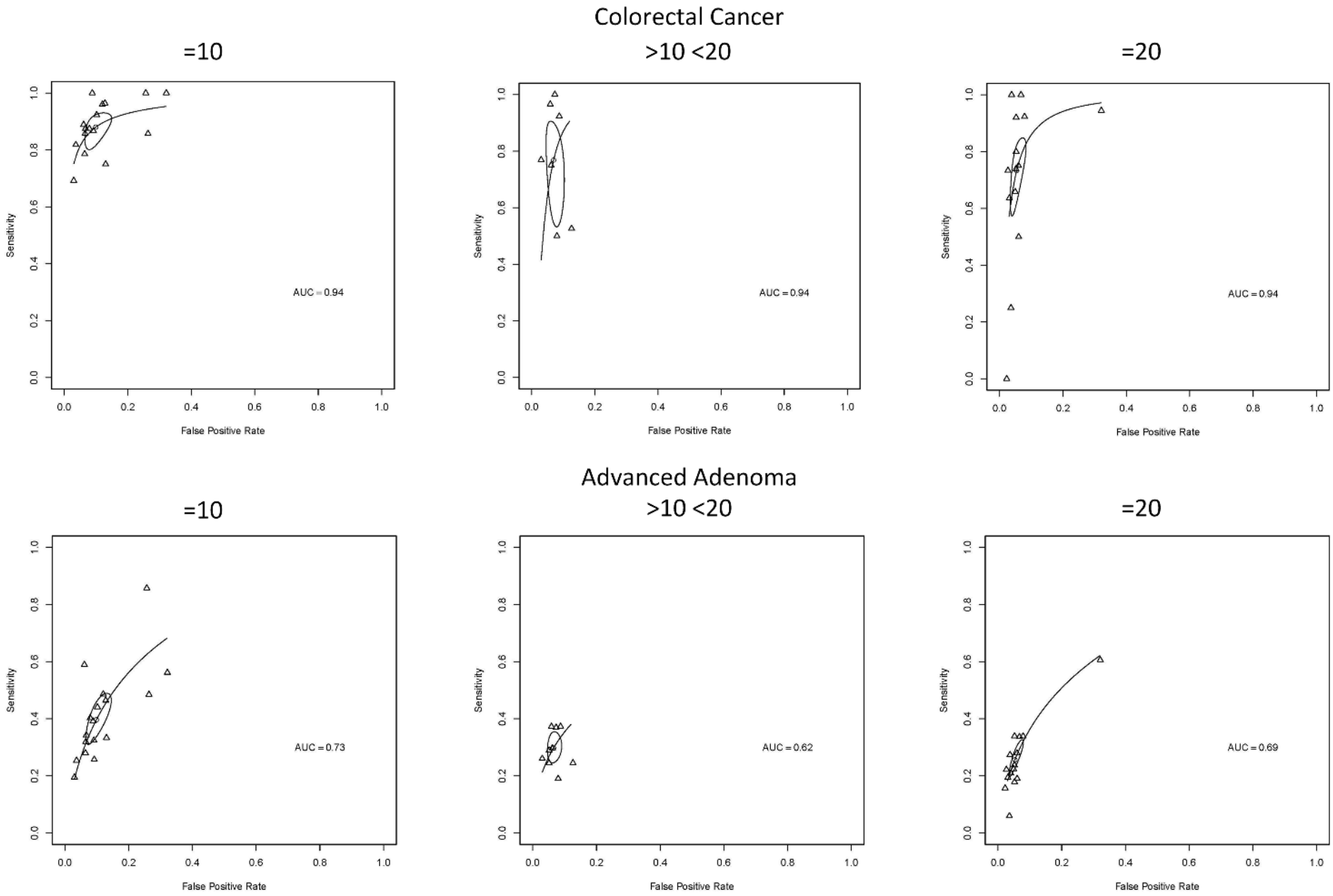
Appendix Figure 1. QUADAS-2 Study Quality Stacked Bar Charts





Q1. Was a consecutive or random sample of patients enrolled?
Q2. Was a case-control design avoided?
Q3. Did the study avoid inappropriate exclusions?
Q4. Could the selection of patients have introduced bias?
Q5. Are there concerns that the included patients do not match the review question?
Q6. Were the index test results interpreted without knowledge of the results of the reference standard?
Q7. If a threshold was used, was it pre-specified?
Q8. Could the conduct or interpretation of the index test have introduced bias?
Q9. Are there concerns that the index test, its conduct, or interpretation differ from the review question?
Q10. Is the reference standard likely to correctly classify the target condition?
Q11. Were the reference standard results interpreted without knowledge of the results of the index test?
Q12. Could the reference standard, its conduct, or its interpretation have introduced bias?
Q13. Are there concerns that the target condition as defined by the reference standard does not match the review question?
Q14. Was there an appropriate interval between index test(s) and reference standard?
Q15. Did all patients receive a reference standard?
Q16. Did all patients receive the same reference standard?
Q17. Were all patients included in the analysis?
Q18. Could the patient flow have introduced bias?

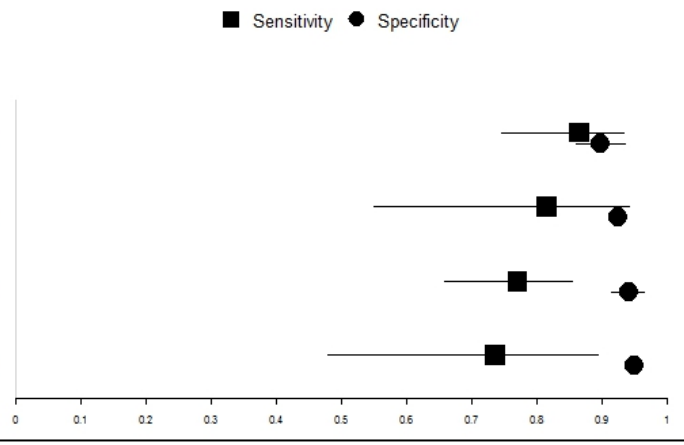
Appendix Figure, 2. Receiver Operating Characteristic Curves by FIT Threshold



Appendix Figure 3a. OC Sensor Summary-level Test Characteristics by Threshold for Colorectal Cancer

Colorectal Cancer - OC Sensor

FIT Threshold µg/g	Study N	Patient N	Patients with CRC	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<= 10	7	6715	56	0.86 (0.75, 0.93), 0.90 (0.86, 0.93)	8.45 (5.36, 12.54), 0.15 (0.07, 0.30)
>10 <20	4	3890	34	0.81 (0.55, 0.94), 0.93 (0.91, 0.93)	10.93 (6.46, 14.45), 0.20 (0.06, 0.49)
=20	11	27828	163	0.77 (0.66, 0.85), 0.94 (0.91, 0.96)	13.88 (7.68, 22.35), 0.28 (0.11, 0.56)
>20	7	4347	41	0.73 (0.48, 0.89), 0.95 (0.94, 0.96)	14.71 (7.68, 22.35), 0.28 (0.11, 0.56)

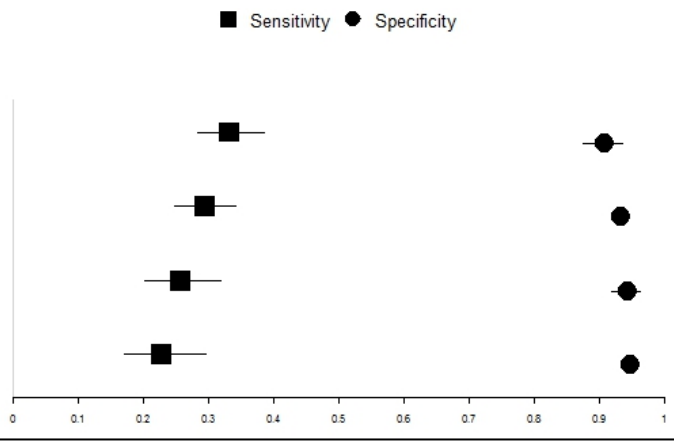


CI: 95% Confidence Interval; CRC: Colorectal Cancer; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 3b. OC Sensor Summary-level Test Characteristics by Threshold for Advanced Adenoma

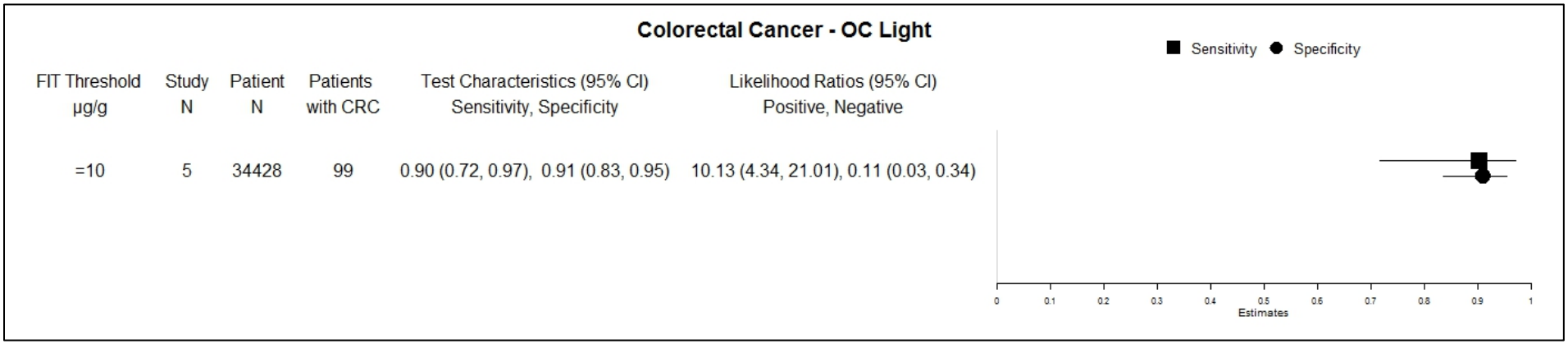
Advanced Adenoma - OC Sensor

FIT Threshold µg/g	Study N	Patient N	Patients with AA	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<=10	9	13153	936	0.33 (0.28, 0.39), 0.91 (0.87, 0.94)	3.66 (2.26, 5.93), 0.73 (0.66, 0.82)
>10 <20	6	10328	702	0.29 (0.25, 0.34), 0.93 (0.92, 0.94)	4.41 (3.16, 6.09), 0.73 (0.66, 0.82)
=20	12	33937	2286	0.26 (0.20, 0.32), 0.95 (0.92, 0.96)	4.70 (2.50, 8.74), 0.79 (0.71, 0.87)
>20	7	4347	354	0.23 (0.17, 0.30), 0.95 (0.94, 0.96)	4.46 (2.72, 7.17), 0.81 (0.73, 0.88)



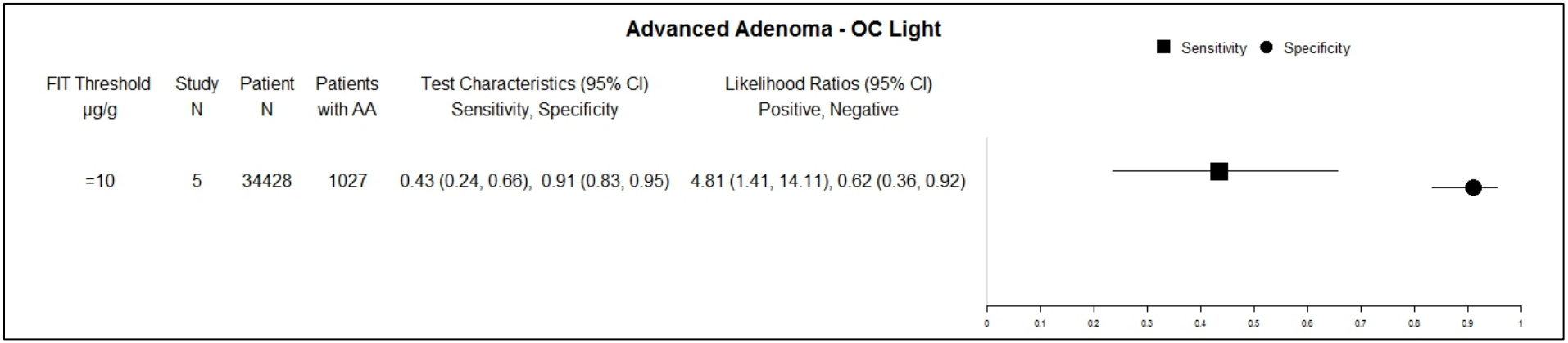
AA: Advanced Adenoma; CI: 95% Confidence Interval; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 4a. OC Light Summary-level Test Characteristics by Threshold for Colorectal Cancer



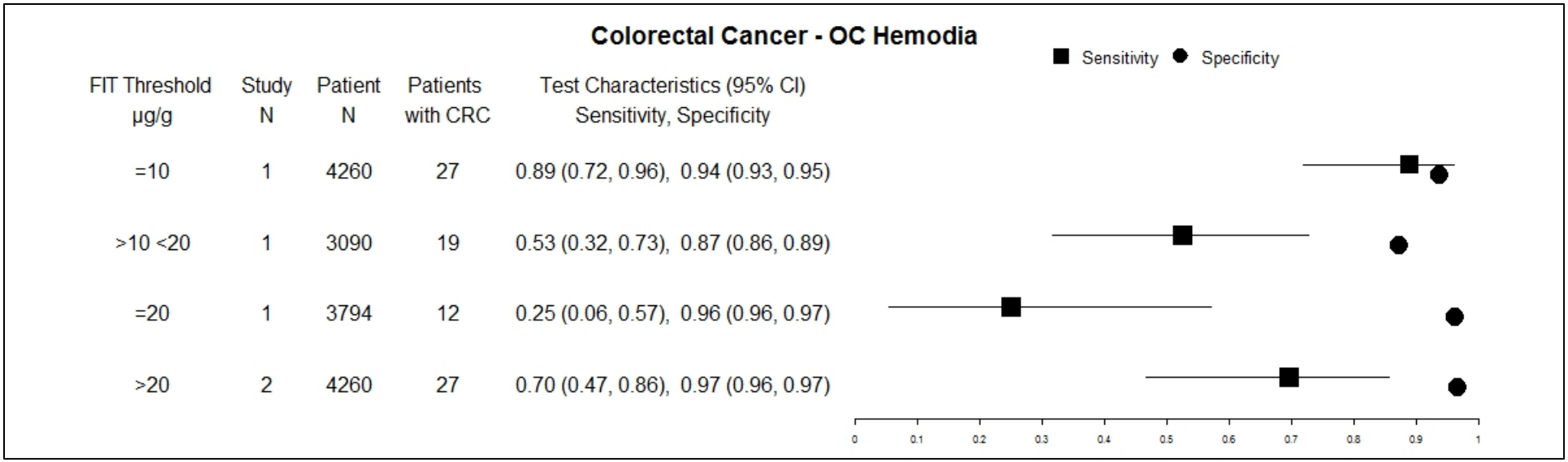
CI: 95% Confidence Interval; CRC: Colorectal Cancer; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 4b. OC Light Summary-level Test Characteristics by Threshold for Advanced Adenoma



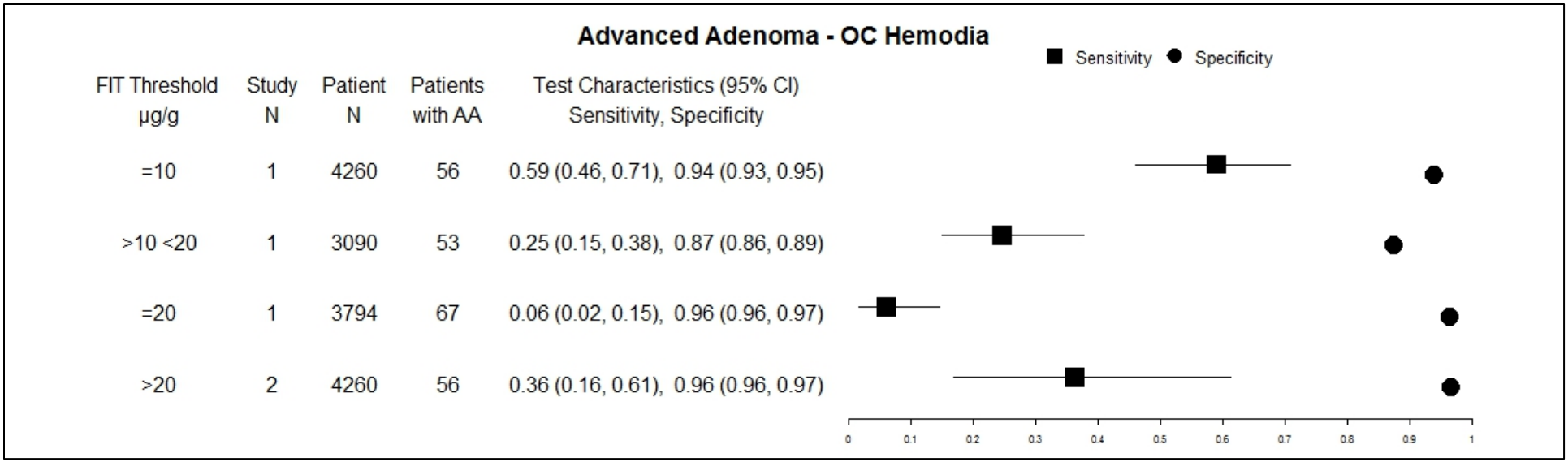
AA: Advanced Adenoma; CI: 95% Confidence Interval; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 5a. OC Hemodia Summary-level Test Characteristics by Threshold for Colorectal Cancer



CI: 95% Confidence Interval; CRC: Colorectal Cancer; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 5b. OC Hemodia Summary-level Test Characteristics by Threshold for Advanced Adenoma

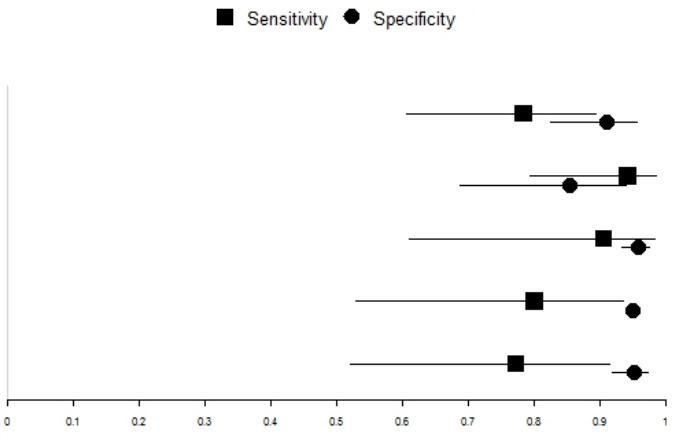


AA: Advanced Adenoma; CI: 95% Confidence Interval; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 6a. Summary-level Test Characteristics by Threshold for Other FITs (Colorectal Cancer)

Colorectal Cancer - Other FIT Tests

FIT Threshold µg/g	Study N	Patient N	Patients with CRC	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<10	8	7279	57	0.78 (0.61, 0.89), 0.91 (0.83, 0.96)	8.78 (3.47, 20.66), 0.24 (0.11, 0.48)
=10	4	5489	43	0.94 (0.79, 0.99), 0.85 (0.69, 0.94)	6.42 (2.54, 16.10), 0.07 (0.02, 0.30)
>10 <20	2	5747	42	0.91 (0.61, 0.98), 0.96 (0.93, 0.97)	21.41 (9.02, 37.53), 0.10 (0.02, 0.42)
=20	2	25016	104	0.80 (0.53, 0.93), 0.95 (0.95, 0.95)	15.91 (9.66, 20.25), 0.21 (0.07, 0.50)
>20	3	8734	49	0.77 (0.52, 0.91), 0.95 (0.92, 0.97)	16.25 (6.29, 34.00), 0.24 (0.09, 0.52)

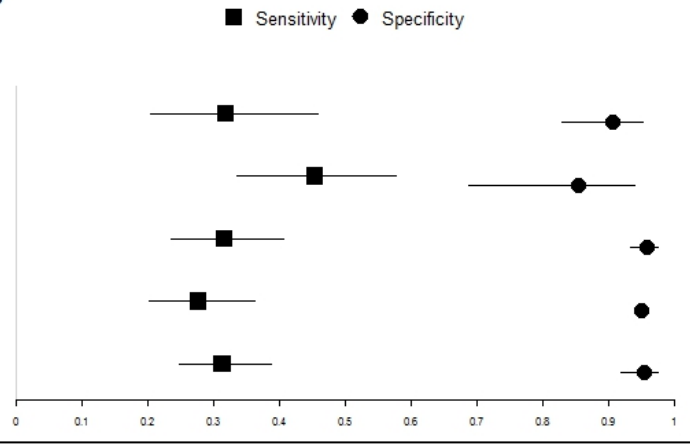


CI: 95% Confidence Interval; CRC: Colorectal Cancer; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model

Appendix Figure 6b. Summary-level Test Characteristics by Threshold for Other FITs (Advanced Adenoma)

Advanced Adenoma - Other FIT Tests

FIT Threshold µg/g	Study N	Patient N	Patients with AA	Test Characteristics (95% CI) Sensitivity, Specificity	Likelihood Ratios (95% CI) Positive, Negative
<10	9	7563	668	0.32 (0.20, 0.46), 0.91 (0.83, 0.95)	3.41 (1.20, 9.36), 0.75 (0.57, 0.96)
=10	4	5489	515	0.45 (0.34, 0.58), 0.85 (0.69, 0.94)	3.11 (1.07, 9.59), 0.64 (0.45, 0.97)
>10 <20	2	5747	569	0.31 (0.24, 0.41), 0.96 (0.93, 0.97)	7.46 (3.49, 15.57), 0.72 (0.61, 0.82)
=20	2	25016	934	0.28 (0.24, 0.41), 0.96 (0.93, 0.97)	5.47 (3.70, 7.84), 0.76 (0.67, 0.84)
>20	3	8734	556	0.31 (0.25, 0.39), 0.95 (0.92, 0.97)	6.64 (2.99, 14.66), 0.72 (0.63, 0.82)



AA: Advanced Adenoma; CI: 95% Confidence Interval; µg/g: microgram of hemoglobin per gram of stool
All results generated using a bivariate model