Erasmus MC
Universitair Medisch Centrum Rotterdam

A quarter of participants with advanced neoplasia have discordant results from 2-sample fecal immunochemical tests for colorectal cancer screening

E. Wieten, C.M. de Klerk, I. Lansdorp-Vogelaar, P.M. Bossuyt, E. Dekker, M.C.W. Spaander

ABSTRACT

Background

Some colorectal cancer (CRC) screening programs use 2-sample fecal immunochemical tests (FITs). We aimed to assess advanced neoplasia (AN) yield of 2 different FIT assays performed on the same bowel movement and have discordant results.

Methods

We conducted a large prospective comparative accuracy study within the Dutch national CRC screening program to evaluate 2 quantitative FIT assays (FOB-Gold, Sentinel, Italy and OC-Sensor, Eiken, Japan) with comparable performance characteristics. We asked 42,179 screening-naïve individuals, 55–75 years old, to perform both tests on the same bowel movement, from May 2016 through March 2017. Participants with ≥1 positive test result (≥15 μg hemoglobin/gram feces) were invited for colonoscopy examination. Detection of AN by colonoscopy was the reference standard.

Results

A total of 21,078 participants (50% participation rate) were included. FIT results were both negative for 19,032 participants (90%), both positive for 1163 participants (5.5%), and discordant for 883 participants (4.2%). AN was detected in 500 participants with 2 positive FIT results (43%) compared to 187 with discordant FIT results (21%) (p<0.001). Of the 687 participants found to have AN by colonoscopy, 187 had only 1 positive FIT result (27%).

Conclusion

In a large 2-sample FIT-based CRC screening study, more than a quarter of participants in whom AN was detected by colonoscopy in the first screening round had discordant FIT results. AN was detected in one-fifth of those with FIT discordance. Participants with discordant results from 2 FITs should undergo colonoscopy. (www.trialregister.nl; no. NTR5874).

INTRODUCTION

Most of the current colorectal cancer (CRC) population-based screening programs using fecal immunochemical tests (FITs) are based on one-sample FIT.[1] In some CRC screening programs, two-sample FITs are used.[1] In clinical practice, screening participants with a positive FIT result sometimes request performing a second FIT (to validate or invalidate the first result), or that two FITs are administered at a short interval. The data addressing CRC screening with more than one FIT are scarce, and most of the evidence is derived from symptomatic patients [2, 3]. When two-sample FITs are used, the follow-up is clear for those with concordant results, but there is uncertainty regarding the appropriate approach to those with discordant results.

In a recently published trial we asked FIT screening participants to perform two tests (FOB-Gold and OC-Sensor) on the same bowel movement.[4] Although both tests were shown to be equivalent for the detection of advanced neoplasia (AN), a proportion of participants had discordant FIT results (one positive and one negative) and some of those participants had AN was detected at colonoscopy.

In this study, we assessed the colonoscopic detection rate of AN in participants with discordant FIT results in a large paired accuracy study conducted within the Dutch nationwide CRC screening program.

METHODS

Study population

The study was conducted within the Dutch CRC screening program between May 2016 and March 2017. Details of this accuracy trial have been provided in detail elsewhere.[4] In short, data of all individuals (55-75 years old), living in the southwest region of the Netherlands, were obtained from municipal registers. A random sample was taken from this target population with a computer-run algorithm (SPSS, IBM, version 23.0). Selection of study invitees preceded invitation. The study was approved by the Dutch Minister of Health (Population Screening Act: publication no. 769500-1357 16-PG) and registered at the Dutch National Trial Registry (no. NTR5874). All study participants gave written informed consent. All authors had access to the study data and approved the final manuscript.

Study design and intervention

Two different FIT assays (one FOB-Gold, Sentinel Diagnostics, Italy and one OC-Sensor, Eiken Chemical, Japan) were sent by postal mail to screening-naïve individuals, who were asked to perform both tests on the same bowel movement. Sample collection instructions were given as recommended by the manufacturers of the tests. Quantitative results for the two different FIT assays were provided in ng hemoglobin (Hb)/mL feces. FOB-Gold

was considered positive at cut-off ≥88 ng/mL and OC-Sensor at cut-off ≥75 ng/mL feces. Converted into micrograms (µg) of Hb per gram of stool, the cut-off for a positive test result was ≥15 µg Hb/g feces for both tests.[5] Participants with either a positive FOB-Gold or positive OC-Sensor result were invited for a pre-colonoscopy interview in an accredited colonoscopy center near the participant's home address.[4] At this pre-colonoscopy interview, participants' eligibility for colonoscopy was assessed. Colonoscopy exclusion criteria were: a life expectancy of five years or less, history of proctocolectomy, undergoing current treatment for CRC, history of inflammatory bowel disease, and a complete colonoscopy in the past five years.[4, 6] For the purpose of the present study, participants were included in the analysis if they had returned two complete tests. A complete test was defined as a returned and analyzable FIT, with a reliable test result. All study invitees received a questionnaire on data regarding body mass index (BMI), use of alcohol, smoking status and use of anticoagulants.

Definitions

Concordant positive and concordant negative FIT results referred to both FOB-Gold and OC-Sensor being positive or negative in the same study participant, respectively. Discordant FIT result referred to a difference in FIT result between the FOB-Gold and OC-Sensor at a positivity cut-off of ≥15 µg Hb/g feces in the same study participant. We regarded participants with discordant FIT results as one group instead of two groups (one positive FOB-Gold result and one negative OC-Sensor and visa versa), as we showed in our previous work that both FITs are equivalent for the detection of AN.[4] The AN detection rate was defined as the number of participants with AN relative to the total number of participants with two complete FIT results (one complete FOB-Gold and one complete OC-Sensor). AN includes CRC and advanced adenoma. An advanced adenoma was defined as an adenoma ≥ 10 mm, with ≥ 25% villous component and/or high-grade dysplasia. When multiple lesions were present in one participant, the participant was classified according to the most advanced lesion.[4] Socioeconomic status (SES) was assessed by the area social status score (combining education, income and employment status) developed by the Netherlands Institute of Social Research, and grouped into quintiles, with 1 being the highest status and 5 the lowest.[7]

Statistical analysis

The median fecal Hb concentration was compared between participants with concordant positive FIT results and those with discordant FIT results using the Mann-Whitney U test. For comparing the most advanced lesion found at colonoscopy between subgroups a chi-square test was used. Positive predictive values are reported with 95% CI.

First, we compared lesion characteristics between those with concordant positive and discordant FIT results. We also compared participant characteristics between three groups: those with concordant positive, concordant negative and discordant FIT results. The Kruskall-Wallis test statistic was used to compare age, BMI, alcohol intake and pack-years of smoking, and chi-square test statistics were used to compare sex, socioeconomic status, smoking status, ethnicity and the use of anticoagulants.

In a secondary analysis, to assess features potentially associated with discordant FIT results, we compared participant characteristics between those with concordant positive FIT results and those with discordant FIT results using multivariable logistic regression analysis. . First, single-variable logistic regression analyses were performed to assess if participant characteristics (sex, age, SES, BMI, alcohol intake, smoking status, ethnicity, and use of anticoagulants) were associated with having discordant test results. Subsequently, all variables with a significant association (p<0.10) were included in a multivariable logistic regression model. Interactions between all variables included in the multivariable model were tested for statistical significance and included in the final model when p<0.05. Hosmer-Lemeshow statistics were used to evaluate goodness-of-fit. The area under the receiving-operator curve was used as a measure of performance of the final model.
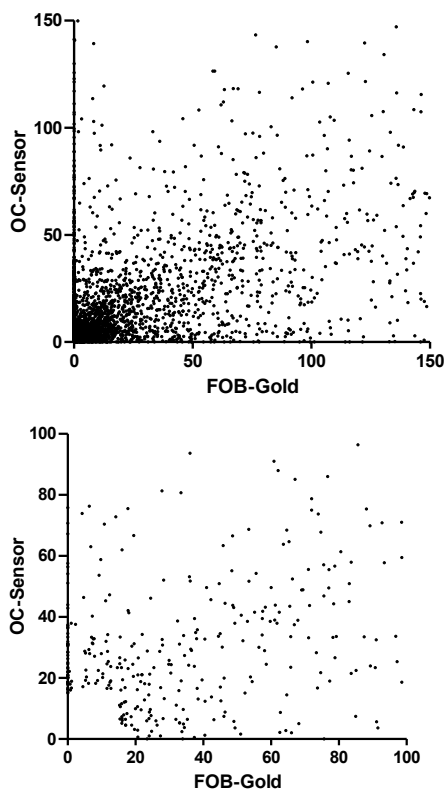
RESULTS

Study cohort

Of the 42,179 screening invitees, 22,111 (52%) returned at least one test and 21,078 (50%) returned both FOB-Gold and OC-sensor. Of the participants who returned both FIT assays, 10,589 (50%) were male. In total, 2,046 (10%) were referred for colonoscopy (one or two positive FIT results).

Figure 1 shows a scatterplot of fecal hemoglobin concentration in μg Hb/g feces as reported with FOB-Gold and OC-Sensor in all participants who returned two complete tests (Figure 1a) and in those in whom AN was detected (true positives for AN, Figure 1b). Figure 1 shows that there is a wide variability in detected hemoglobin by the two FITs, even when taken from the same fecal sample. In the 1724 (84%) participants attending colonoscopy, AN was detected in 687 (40%) and CRC in 80 (4.6%). The overall positive predictive value (PPV) for AN was 34%, and 4% for CRC.

Concordant and discordant FIT results and detection of advanced neoplasia

Of the 21,078 (50%) participants who returned two complete tests, FIT results (FOB-Gold and OC-Sensor) were concordant positive (at a positivity cut-off of 15 μg Hb/g feces) in 1,163 (5.5%) participants, concordant negative in 19,032 (90%), and discordant in 883 (4.2%).

Erasmus University Rotterdam

Figure 1 Scatterplot of FOB-Gold versus OC-Sensor in fecal hemoglobin in µg Hb/gram feces detected in a) all participants with two complete tests and b) in true positives for AN..

In participants with concordant positive test results, significantly higher fecal Hb concentrations were observed than in participants with discordant test results (Table 1). AN was found at colonoscopy in 500 of 1163 (43%) participants with two positive FITs, compared to 187 of 883 (21%) participants with discordant FIT results (p<0.001; Table 1).

CRC was detected in 70 (7.2%) of those with two positive FIT results compared to 10 (1.3%) in those with discordant FIT results (p<0.001; Table 1). Supplementary Table 1 shows AN and CRC detection rates for the same subgroups at higher positivity cut-offs.

Of the 687 participants in whom advanced neoplasia was detected, 187 (27%) had discordant results of FOB-Gold and OC-Sensor.

Characteristics of participants with concordant and discordant FIT results
Table 2 shows participant characteristics for those with concordant positive, discordant, and concordant negative FIT results. When comparing these three groups, significant

Table 1 Detection of fecal Hb concentration, most advanced lesion and associated positive predictive values of AN and CRC in study participants with concordant and discordant FIT results (FOB-Gold and OC-Sensor)*

	Concordant positive FIT results	Discordant FIT results	P-value
Fecal Hb concentration median, (IQR)			
Complete tests	n=1163	n=883	
OC-Sensor	63 (32-182)	16 (6-28)	<0.01
FOB-Gold	110 (51-220)	13 (0-33)	<0.01
Most advanced lesion (n, %)			
Advanced neoplasia	500 (43)	187 (21)	<0.01
CRC	70 (6)	10 (1)	<0.01
Advanced adenoma	430 (37)	177 (20)	<0.01
Other malignancy	1 (0.1)	0	0.38
Non-advanced adenoma**	242 (21)	236 (28)	<0.01
Serrated lesion	41 (4)	66 (8)	<0.01
Positive predictive value, % (95%CI)			
Advanced neoplasia	43 (40-46)	21 (19-24)	
CRC	6 (5-8)	1 (1-2)	

*at a positivity cutoff level of 15 µg Hb/g feces
** 73 cases missing for non-advanced adenomas yes/no
AN: advanced neoplasia; CRC: colorectal cancer; FIT: fecal immunochemical test

differences were found for sex, age, SES, BMI, weekly alcohol intake, smoking status, pack-years of smoking, self-reported ethnicity, and use of anticoagulants.

Logistic regression analyses were performed to identify variables associated with FIT discordance (Table 3). A lower BMI (p=0.03) and lower median number of pack-years of smoking (p=0.005) were significantly associated with discordant FIT results in both univariable and multivariable regression analyses, with an odds ratio of 0.972 per kg/m^2 increase (p=0.009) and 0.992 per pack-year of smoking increase (p=0.008) in the multivariable model, respectively. There were no significant interactions between variables. The Hosmer-Lemeshow test goodness-of-fit test of the multivariable model was not significant (p=0.61). The area under the receiving-operator curve of this final model was 0.558 (p<0.001).

DISCUSSION

In this large prospective study, performed within the Dutch nationwide CRC screening program, the proportion of participants with discordant FIT results almost equaled the proportion of those with two positive FIT results, at a cut-off of 15 µg Hb/gram feces. In one fifth of those with discordant FIT results, advanced neoplasia was detected. These findings imply the following: (1) colonoscopy is indicated in case a screenee has one positive and

Erasmus University Rotterdam

Table 2 Concordant and discordant FIT results of FOB-Gold and OC-Sensor and participant characteristics

	Concordant positive FIT results n=1163	Discordant FIT results n=883	Concordant negative FIT results n=19,032	P-value
Participant characteristics				
Male n, (%)	701 (60)	518 (59)	9370 (49)	<0.01
Age median (IQR)	61 (59-70)	60 (59-63)	60 (59-62)	<0.01
Socio-economic status n, (%)				<0.01
Very high	198 (17)	188 (21)	3840 (20)	
High	218 (19)	159 (18)	4275 (22)	
Average	208 (18)	153 (17)	3515 (19)	
Low	310 (27)	207 (23)	4402 (23)	
Very low	227 (20)	174 (20)	2957 (15)	
Missing	2 (0.2)	2 (0.2)	43 (0.2)	
BMI^ in kg/m^2 median, (IQR)	27 (24-30)	26 (24-29)	26 (24-29)	<0.01
Alcohol intake^ glasses per week* mean, (IQR)	8 (3-15)	7 (3-14)	6 (2-10)	<0.01
Smoking status n, (%)				<0.01
Never smoker	283 (24)	240 (27)	6766 (36)	
Former smoker	570 (49)	439 (50)	8757 (46)	
Current smoker	275 (24)	180 (20)	2994 (16)	
Missing	35 (3)	24 (3)	514 (3)	
Pack-years of smoking				
All participants^ median, (IQR)	1.5 (0-20)	1 (0-17)	0 (0-10)	
(Ever) smoker^ median, (IQR)	20 (8-34)	15 (5-30)	12 (4-25)	
Ethnicity** n, (%)				<0.01
Dutch	1066 (92)	816 (92)	17,502 (92)	
Turkish or Moroccan	4 (0.3)	2 (0.2)	52 (0.3)	
Hindustan	6 (0.5)	4 (0.5)	81 (0.4)	
Surinamese or Creole	2 (0.2)	1 (0.1)	64 (0.3)	
Mix of ethnicities	26 (2)	11 (1)	367 (2)	
Other	31 (3)	24 (3)	472 (3)	
Missing	28 (2)	25 (3)	494 (3)	
Use of anticoagulants*** ^ n, (%)	327 (28)	243 (28)	4330 (23)	<0.01

*In the year preceding the questionnaire

**Self-reported ethnicity by the participant in the questionnaire

*** Use of any of the following anticoagulants ≥ 3 per week in one month preceding the questionnaire: acetylsalicylic acid,carbasalate calcium, diclofenac, ibuprofen, naproxen, etoricobix, meloxicam, diclofenac/misoprotol, coumarin derivate or combinations

^ Complete cases from questionnaires for BMI n=20,250, alcohol intake n=15,650, pack-years of smoking n=17,792, and use of anticoagulants n=20,727

Table 3 Univariable and multivariable regression analyses of participant characteristics associated with discordant FIT results versus concordant positive FIT results

	Univariable OR (95% CI)	P-value	Multivariable OR (95% CI)	P-value
Participant characteristics				
Male n, (%)	0.94 (0.78-1.12)	0.46	)	
Age per year increase	0.99 (0.98-1.01)	0.46		
Socio-economic status n, (%)		0.13		
Very high	*Reference*			
High	0.87 (0.67-1.14)			
Average	0.96 (0.72-1.28)			
Low	0.95 (0.72-1.26)			
Very low	1.24 (0.94-1.64)			
BMI per kg/m^2 increase	0.98 (0.96-1.00)	0.03	0.97 (0.95-0.99)	0.01
Alcohol per weekly increase*, (IQR)	0.99 (0.99-1.00)	0.18		
Smoking status n, (%)		0.13		
Never smoker	*Reference*			
Former smoker	0.91 (0.73-1.12)			
Current smoker	0.77 (0.60 -1.00)			
Pack-years of smoking per year increase	0.99 (0.99-1.00)	0.03	0.99 (0.99-1.00)	0.01
Ethnicity** n, (%)		0.66		
Mix of ethnicities	*Reference*			
Dutch	1.81 (0.89-3.68)			
Turkish or Moroccan	1.18 (0.19-7.43)			
Hindustan	1.58 (0.37-6.71)			
Surinamese or Creools	1.18 (0.097-14.42)			
Other	1.83 (0.76-4.43)			
Use of anticoagulants** n, (%)	1.03 (0.85-1.25)	0.77		

*In the year preceding the questionnaire

**Self-reported ethnicity by the participant in the questionnaire

*** Use of any of the following anticoagulants ≥ 3 per week in one month preceding the questionnaire: acetylsalicylic acid, carbasalate calcium, diclofenac, ibuprofen, naproxen, etoricobix, meloxicam, diclofenac/misoprotol, coumarin derivate or combinations

one negative FIT, and (2) offering a second confirmatory FIT to screenees with a positive FIT should be discouraged, given the high rate of advanced neoplasia detected in this group even at a generally low positivity cut-off.

Our study findings do not support two-sample FIT screening to improve AN detection rate over other FIT screening strategies, such as the use of a lower positivity cut-off, as most AN were detected by both tests. We did not find a correlation between participant

Erasmus University Rotterdam

characteristics and a discordant FIT result except for lower pack-years of smoking and BMI, although the clinical utility of these findings is uncertain, especially in light of the poor performance of the multivariable model. Known risk factors for CRC (male, higher age, low SES, high BMI, alcohol intake, pack-years of smoking) were all significantly associated with having two positive FIT results, indicating that a higher pre-test probability of having AN is associated with a higher probability of being detected by both FITs.

Obtaining discordant FIT results could be attributed to sampling issues, because of nonhomogeneous distribution of hemoglobin through the feces. Alternatively, having two different FIT results may also be due to differences between the two tests (FOB-Gold and OC-Sensor). These tests have different self-sampling instructions, sample collection devices, buffer volumes, antibodies against different globin epitopes, and analytical systems. These differences may all contribute to variability in the level of fecal hemoglobin detected (Figure 1).[8, 9] Recently, it was suggested that for a fair comparison of FIT results from different brands, similar positivity rates should be used rather than similar positivity cut-offs.[10] Our findings suggest that discordant FIT results are less likely attributable to test differences, but rather to nonhomogeneous distribution of hemoglobin through the feces. First, we have previously reported that AN detection rates are equivalent between FOB-Gold and OC-Sensor at the same cut-off of 15 µg Hb/g feces.[4] Second, another smaller two-sample study, of two FITs of the same brand performed on consecutive days reported results comparable to ours: the positivity rates were 8.4% and 12.7% and positive predictive values for AN were 40% and 34%, for one-sample and two-sample FIT screening in first round invitees, respectively.[11]

We found that detection rates for non-advanced adenomas and serrated lesions were higher in participants with discordant FIT results than in participants with concordant positive FIT results. The higher detection of these lesions is probably due to how these lesions are registered in the Dutch nationwide CRC screening program, where lesions are classified according to the most advanced finding. Non-advanced lesions and serrated lesions were probably reported less often in participants with two positive FIT results because more advanced lesions were detected in this group compared to participants with discordant FIT results. Another contributing cause is the reduced bleeding of these lesions compared to advanced neoplasms.

Some limitations of our study should be acknowledged. First, our results are limited to first round screenees and a pre-set FIT positivity cut-off. Though future research could assess the effects of paired two-sample screening on subsequent rounds, we already found a substantial proportion of participants with a discordant FIT results at this rather low FIT positivity cut-off (15 µg Hb/gram feces for both FITs). There are even higher discordance

rates and missed AN at higher positivity cut-offs, as shown in Supplementary Table 1. Second, we regarded participants with discordant FIT results as one group, not as two groups (positive FOB-Gold result and negative OC-Sensor and visa versa), because the two FITs have been shown to be equivalent for the detection of advanced neoplasia.[4] Finally, most participant characteristics were based on a questionnaire and not additionally verified. However, the questionnaire had been tested and validated in a previous study.[12]

In conclusion, this large prospective paired FIT study has highlighted that a substantial proportion of advanced neoplasia was detected in screenees with discordant FIT results. In case a participant has one positive and one negative FIT result, colonoscopy should strongly be advised.

REFERENCES

1. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637–49.

2. Auge JM, Fraser CG, Rodriguez C, Roset A, Lopez-Ceron M, Grau J, et al. Clinical utility of one versus two faecal immunochemical test samples in the detection of advanced colorectal neoplasia in symptomatic patients. Clin Chem Lab Med. 2016;54(1):125-32.

3. Hogberg C, Soderstrom L, Lilja M. Faecal immunochemical tests for the diagnosis of symptomatic colorectal cancer in primary care: the benefit of more than one sample. Scand J Prim Health Care. 2017;35(4):369-72.

4. Wieten E, de Klerk CM, van der Steen A, Ramakers CR, Kuipers EJ, Hansen BE, et al. Equivalent Accuracy of 2 Quantitative Fecal Immunochemical Tests in Detecting Advanced Neoplasia in an Organized Colorectal Cancer Screening Program. Gastroenterology. 2018;155(5):1392-9 e5.

5. Fraser CG, Allison JE, Halloran SP, Young GP. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. J Natl Cancer Inst. 2012;104(11):810-4.

6. Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-off Levels. Gastroenterology. 2016;152(4):767-75.

7. Netherlands Institute of Social Research. Socioeconomic status scores The Netherlands. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores (date accessed 20 March 2017); 2016.

8. Rapi S, Berardi M, Cellai F, Ciattini S, Chelazzi L, Ognibene A, et al. Effects of fecal sampling on preanalytical and analytical phases in quantitative fecal immunochemical tests for hemoglobin. Int J Biol Markers. 2017;32(3):e261-e6.

9. Fraser CG. Comparison of quantitative faecal immunochemical tests for haemoglobin (FIT) for asymptomatic population screening. Translational Cancer Research. 2016:S916-S9.

10. Grobbee EJ, van der Vlugt M, van Vuuren AJ, Stroobants AK, Mundt MW, Spijker WJ, et al. A randomised comparison of two faecal immunochemical tests in population-based colorectal cancer screening. Gut. 2017;66(11):1975-82.

11. Schreuders EH, Grobbee EJ, Nieuwenburg SAV, Kapidzic A, van Roon AHC, van Vuuren AJ, et al. Multiple rounds of one sample versus two sample faecal immunochemical test-based colorectal cancer screening: a population-based study. Lancet Gastroenterol Hepatol. 2019.

12. de Wijkerslooth TR, de Haan MC, Stoop EM, Bossuyt PM, Thomeer M, Essink-Bot ML, et al. Burden of colonoscopy compared to non-cathartic CT-colonography in a colorectal cancer screening programme: randomised controlled trial. Gut. 2012;61(11):1552-9.

Supplementary Table 1 Detection of AN and CRC in study participants with concordant and discordant FIT results (FOB-Gold and OC-Sensor)* at different positivity cut-offs in μg Hb/gram feces.

	Concordant positive FIT results	Discordant FIT results	P-value
At positivity cut-off 20	n=1033	n=776	
AN n (%)	462 (45)	177 (23)	<0.001
CRC n (%)	68 (7)	10 (1)	<0.001
At positivity cut-off 30	n=874	n=782	
AN n (%)	419 (48)	180 (23)	<0.001
CRC n (%)	66 (8)	8 (1)	<0.001
At positivity cut-off 40	n=754	n=820	
AN n (%)	375 (50)	200 (24)	<0.001
CRC n (%)	61 (8)	11 (1)	<0.001
At positivity cut-off 50	n=654	n=891	
AN n (%)	345 (53)	222 (25)	<0.001
CRC n (%)	60 (9)	11 (1)	<0.001

AN: advanced neoplasia; CRC: colorectal cancer; Hb: hemoglobin; FIT: fecal immunochemical test

Erasmus University Rotterdam