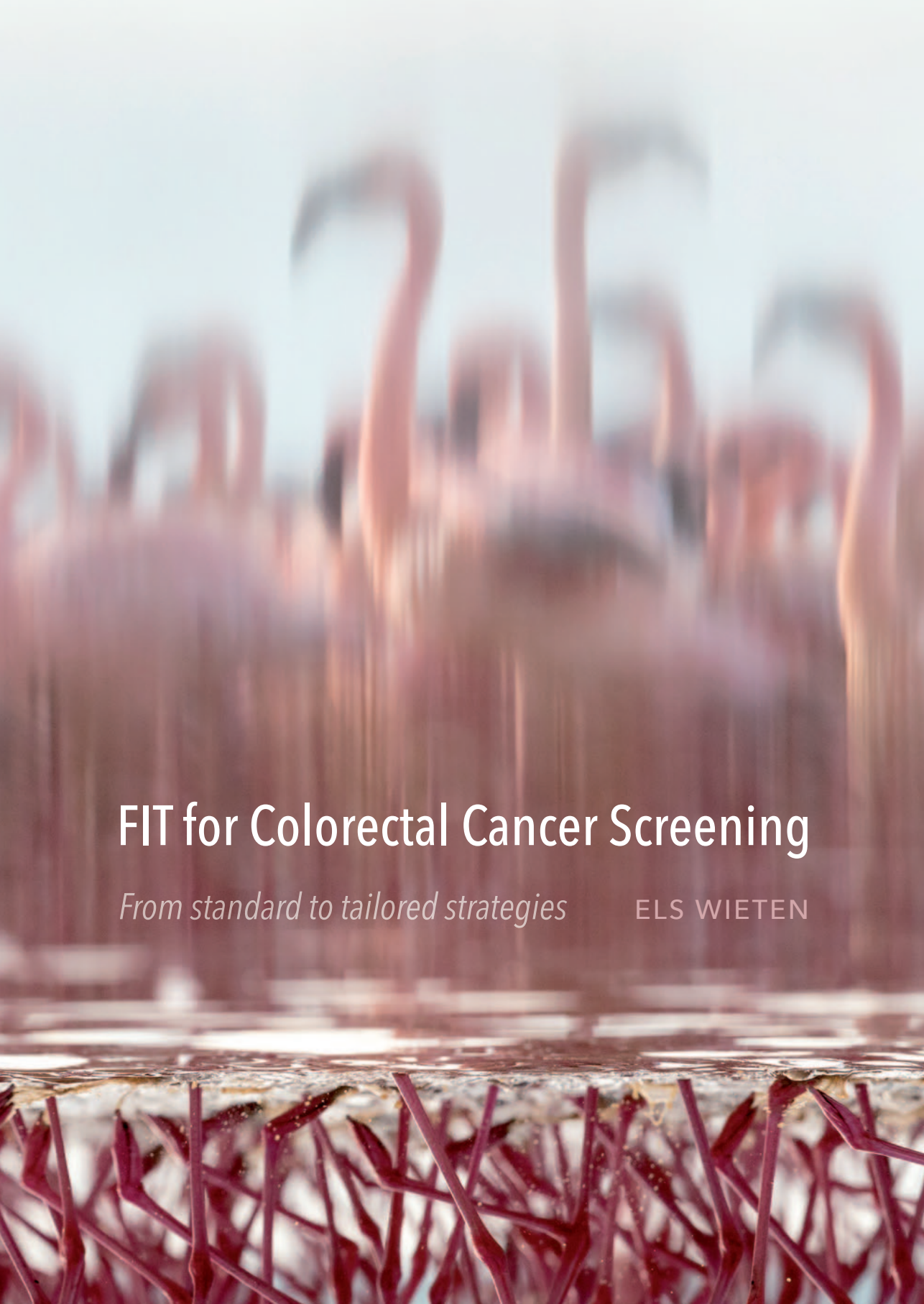


FIT for Colorectal Cancer Screening

From standard to tailored strategies

ELS WIETEN

FIT for Colorectal Cancer Screening
from standard to tailored strategies

Els Wieten

Colophon

The work presented in this thesis was conducted at the department of Gastroenterology and Hepatology, Erasmus MC Unviversity Medical Center, Rotterdam, The Netherlands.

ISBN: 978-94-6361-343-9

Cover photograph by Paul McKenzie
Cover design and printing by Optima Grafische Communicatie

Financial support for printing this thesis was kindly given by the Department of Gastroenterology and Hepatology of Erasmus MC Rotterdam, Erasmus University Rotterdam, Nederlandse Vereniging voor Gastroenterologie, Daklapack, Bevolkingsonderzoek Zuid-West, Post NL, Topicus, Rabobank, ABN AMRO, Norgine, Dr. Falk Pharma Benelux, ChipSoft, Castor EDC, Tramedico, Ferring Farmaceuticals, Pentax Medical and Pfizer.

FIT for Colorectal Cancer Screening
from standard to tailored strategies

FIT voor een bevolkingsonderzoek darmkanker
van standaard naar gepersonaliseerde strategieën

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof. dr. R.C.M.E. Engels
en volgens het besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 18 december 2019 om 9.30 uur

Els Wieten

geboren op woensdag 17 december 1986
te Dordrecht

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor

Prof. dr. M.J. Bruno

Overige leden

Prof. dr. E. Dekker

Dr. I. Lansdorp-Vogelaar

Prof. dr. C. Verhoef

Copromotor

Dr. M.C.W. Spaander

CONTENTS

1

Introduction and outline of the thesis

INTRODUCTION

Colorectal cancer (CRC) is an important health problem; it is the third most common cancer in men and second in women globally.[1] In Europe, CRC is the most common cause of cancer death after lung cancer.[2] Age is a major risk factor for the development of CRC, with the majority of patients with sporadic cancer being > 50 years of age.[3] Other common risk factors include a first-degree relative with CRC and excess body weight.[4, 5]

The lifetime risk of developing colorectal cancer in many regions is around 5%. Over the past years, treatment modalities have largely improved, but still 40-50% of patients with symptomatic CRC eventually die of metastatic disease.[6] These treatment advancements have been accompanied by increased treatment costs. In the United States, colorectal cancer was estimated to be the second highest cancer site with highest national cost of cancer care in 2010, with approximately 14.14 billion dollar.[7] For patients with non-metastasized CRC, surgery is the main curative treatment, which has been associated with appreciable morbidity and mortality rates.[8] Most CRCs develop from an adenoma, the preclinical precursor of CRC.[9] However, only a minority of adenomas ultimately progress to CRC. The transition from early adenoma to invasive colorectal cancer takes years.[10] These characteristics of CRC make CRC particularly suitable for screening.[11, 12]

Screening for colorectal cancer

In recent years, more than 50 countries have implemented population CRC screening.[13] It has been demonstrated that CRC screening reduces both CRC-related mortality and incidence.[12, 14-20] Screening aims to lower the burden of cancer by detecting the disease at an early, preclinical stage.[12, 14, 17, 21]

There are several CRC screening methods available, which vary in the level of supporting evidence, effectiveness, invasiveness, and uptake. Currently CRC screening programs are either based on direct endoscopic visualization of the colon (colonoscopy or flexible sigmoidoscopy) or use fecal occult blood testing (FOBT) as primary screening method. In the latter form of CRC screening, colonoscopy is offered in case of a positive test.

Fecal occult blood testing

Two types of fecal occult blood testing for CRC screening are available: guaiac fecal occult blood testing (gFOBT) and fecal immunochemical testing (FIT). Randomized controlled trials have shown that screening with gFOBT is associated with a 15%-33% decrease in CRC-related mortality rates.[15, 16, 22] Worldwide, FIT is now rapidly replacing gFOBT, as FIT has been shown to be more sensitive for the detection of both CRCs and it's precursors than gFOBT.[23, 24] Fecal immunochemical tests detect human-specific globin of blood, whereas guaiac fecal occult blood tests react with heme, including consumed non-human

heme. Other advantages of FIT over gFOBT include that these tests are easier to handle, which leads to higher participation and allow for single stool testing.[25, 26] They also allow automated handling and may provide quantitative test results. Consequently, the positivity cut-off is adjustable to match available colonoscopy resources.[25, 26]

Due to these advantages of FIT over gFOBT, the European guidelines recommend the quantitative immunochemical tests as test of choice for population CRC screening.[27] More than 50 FIT brands are widely available.[28] However, only few data are available on differences between FIT-brands in screening effectivity to detect advanced neoplasia (AN). FIT brands vary in sampling tubes and buffer volumes, resulting in different fecal hemoglobin measurements that are incomparable.[29, 30]

Current status in The Netherlands

In the Netherlands, a nationwide FIT-based CRC screening program has been gradually implemented from January 2014 onwards. Individuals, aged 55-75 years, are biannually invited to perform a single test, followed by subsequent colonoscopy in case of a positive test. It was decided to start screening with the FOB-Gold (Sentinel, Italy) through a national tender. To match available colonoscopy resources, the positivity cut-off used was increased from 15 to 47 µg hemoglobin per gram feces halfway 2014.[31]

The European guidelines indicate that a screening program should assess individual device characteristics, including accuracy, ease of use by participant and laboratory, suitability for transport, sampling reproducibility and sample stability.[27]

AIMS OF THE THESIS

The aims of this thesis are to compare different fecal occult blood tests for CRC screening and to explore tailored FIT-based screening strategies.

Outlines of the thesis

Interval cancer rate is a key quality indicator in screening programs. Since data on the incidence rate of interval cancers following negative occult stool tests were limited, we performed a systematic literature search and meta-analysis to determine the pooled incidence rates of interval cancers following a negative gFOBT and FIT in **Chapter 2**. In this chapter, we also assessed how these two types of tests compare with regard to interval cancer incidence. In **Chapter 3**, we assess the accuracy of two FIT assays in detecting advanced neoplasia in the Dutch CRC screening program. For this large population-based study, we use a paired design, in which both tests are compared within the same individual and sampled from the same stool. Such design minimizes the risk of confounding factors and increases the applicability of the study results to CRC screening programs worldwide.

To facilitate further informed decision-making on implementing one of both tests for a CRC screening program, we assessed participation rates and ease of use of the tests in **Chapter 4.**

Furthermore, we compare the accuracy of the two FIT assays in detecting advanced neoplasia across various test positivity cut-offs in **Chapter 5**.

Second, we explore tailored CRC screening strategies with FIT. Potential factors of use for tailored or personalized screening may include sex, age, family history, genetic and environmental factors, lifestyle, fecal hemoglobin levels detected (over time) and multiple sample screening. In **Chapter 6**, we assess the diagnostic yield of two-sample screening and provide information for the decision-making on how to deal with two discordant FIT results. We further analyze positivity rates and detection rates of advanced neoplasia across age categories in **Chapter 7**, and assess how these relate to the positive predictive values and colonoscopy demand at multiple test positivity cut-offs. Finally, we illustrate the effect of gender-tailored FIT screening in **Chapter 8.**

The impact of CRC screening as well as the balance between screening burden and benefits strongly depends on the quality of colonoscopy. Besides quality, safety of the endoscopic procedure and patient satisfaction are important outcome parameters for a screening program. In the final chapter of the thesis, **Chapter 9,** we describe the requirements for accrediting screening centers as well as individual endoscopists in a CRC screening program.

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019;144(8):1941-53.
2. Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer. 2018;103:356-87.
3. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Primers. 2015;1:15065.
4. Kharazmi E, Fallah M, Sundquist K, Hemminki K. Familial risk of early and late onset cancer: nationwide prospective cohort study. BMJ. 2012;345:e8076.
5. Brandstedt J, Wangefjord S, Nodin B, Gaber A, Manjer J, Jirstrom K. Gender, anthropometric factors and risk of colorectal cancer with particular reference to tumour location and TNM stage: a cohort study. Biol Sex Differ. 2012;3(1):23.
6. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893-917.
7. Mariotto AB, Yabroff KR, Shao Y, Feuer EJ, Brown ML. Projections of the cost of cancer care in the United States: 2010-2020. J Natl Cancer Inst. 2011;103(2):117-28.
8. Rabeneck L, Paszat LF, Hilsden RJ, Saskin R, Leddin D, Grunfeld E, et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. Gastroenterology. 2008;135(6):1899-906, 906 e1.
9. Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. Natural history of untreated colonic polyps. Gastroenterology. 1987;93(5):1009-13.
10. 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):1585-89.
11. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease] Principios y metodos del examen colectivo para identificar enfermedades. Bol Oficina Sanit Panam. 1968;65(4):281-393.
12. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med. 1993;329(27):1977-81.
13. 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.
14. Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med. 2012;366(8):687-96.
15. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-77.
16. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, 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):1365-71.
17. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev. 2007(1):CD001216.

18. Jorgensen OD, Kronborg O, Fenger C. A randomised study of screening for colorectal cancer using faecal occult blood testing: results after 13 years and seven biennial screening rounds. Gut. 2002;50(1):29-32.
19. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut. 2015;64(5):784-90.
20. Holme O, Loberg M, Kalager M, Bretthauer M, Hernan MA, Aas E, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. JAMA. 2014;312(6):606-15.
21. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375(9726):1624-33.
22. Kronborg O, Fenger C, Olsen J, Jørgensen OD. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348(9040):1467-71.
23. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100(7):1103-10.
24. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327-37.
25. Kuipers EJ, Spaander MC. Colorectal Cancer Screening by Colonoscopy, CT-Colonography, or Fecal Immunochemical Test. J Natl Cancer Inst. 2016;108(2):djv383.
26. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol. 2013;10(3):130-42.
27. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
28. Daly JM, Xu Y, Levy BT. Which Fecal Immunochemical Test Should I Choose? J Prim Care Community Health. 2017;8(4):264-77.
29. 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.
30. Chiang TH, Chuang SL, Chen SL, Chiu HM, Yen AM, Chiu SY, et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cutoff concentration in a nationwide colorectal cancer screening program. Gastroenterology. 2014;147(6):1317-26.
31. 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. 2017;152(4):767-75.e2.

2

Incidence of faecal occult blood test interval cancers in population-based colorectal cancer screening: a systematic review and meta-analysis

E. Wieten, E.H. Schreuders, E.J. Grobbee, D. Nieboer, W.M. Bramer, I. Lansdorp-Vogelaar, M. J. Bruno, E.J. Kuipers, M.C.W. Spaander.

Gut 2019

ABSTRACT

Objective

Faecal immunochemical tests (FITs) are replacing guaiac faecal occult blood tests (gFOBTs) for colorectal cancer (CRC) screening. Incidence of interval colorectal cancer (iCRC) following a negative stool test result is not yet known. We aimed to compare incidence of iCRC following a negative FIT or gFOBT.

Design

We searched Ovid Medline, Embase, Cochrane Library, Science Citation Index, PubMed and Google scholar from inception to December 12th, 2017 for citations related to CRC screening based on stool tests. We included studies on FIT or gFOBT iCRC in average-risk screening populations. Main outcome was pooled incidence rate of iCRCs per 100,000 person-years (p-y). Pooled incidence rates were obtained by fitting random effect Poisson regression models.

Results

We identified 7426 records, and included 29 studies. Meta-analyses comprised data of 6,987,825 subjects with a negative test result, in whom 11,932 screen-detected CRCs and 5548 gFOBT or FIT iCRCs were documented. Median faecal haemoglobin (Hb) positivity cut-off used was 20 (range 10-200) microgram Hb/gram faeces in the 17 studies that provided FIT results. Pooled incidence rates of iCRC following FIT and gFOBT were 20 (95%CI 14-29;I^2=99%) and 34 (95%CI 20-57;I^2=99%) per 100,000 p-y, respectively. Pooled incidence rate ratio of FIT versus gFOBT iCRC was 0.58 (95%CI 0.32-1.07;I^2=99%) and 0.36 (95%CI 0.17-0.75;I^2=10%) in sensitivity analysis. For every FIT iCRC, 2.6 screen-detected CRCs were found (ratio 1:2.6), for gFOBT the ratio between iCRC and screen-detected CRC was 1:1.2. Age below 60 years, and the third screening round were significantly associated with a lower iCRC rate.

Conclusion

A negative gFOBT result is associated with a higher iCRC incidence than a negative FIT. This supports the use of FIT over gFOBT as CRC screening tool.

INTRODUCTION

Worldwide, colorectal cancer (CRC) is the third most common cancer in men and the second in women.[1] Randomised controlled trials have shown that screening with guaiac faecal occult blood tests (gFOBTs), and subsequent colonoscopy if the result is positive, is associated with a 15%-33% decrease in CRC-related mortality.[2-4] Consequently, these stool tests are widely used for CRC screening.[5]

A cost-effective screening program is inevitably associated with the occurrence of what are known as interval colorectal cancers (iCRCs) – defined as CRCs detected after a negative screening test and before the next recommended test is due.[6, 7] The rate of occurrence is strongly related to the sensitivity of a screening test and reflects the quality of a screening program.[8, 9] Therefore, the International Agency for Research on Cancer (IARC) recommends to collect and report data on iCRCs.[8] In CRC screening programs, iCRCs are cases either missed by stool tests or at colonoscopy.[7] Prevalence and associated risk factors of post-colonoscopy iCRCs in screening programs have been described.[10, 11] However, data on prevalence and associated risk factors of iCRCs following negative occult stool tests are still lacking.

In fecal occult blood test (FOBT)-based CRC screening programs, gFOBTs have been the most commonly used occult stool tests for years. At present, gFOBT is rapidly replaced by fecal immunochemical testing (FIT).[5] FIT detects human-specific globin, whereas gFOBTs react with heme, including consumed non-human heme. FITs are more sensitive for the detection of CRC as well as its precursors than gFOBTs.[12, 13] Moreover, FITs allow single stool testing, are easier to handle, have higher participation rates and provide quantitative test results, which enables to adjust the positivity cut-off to match available resources.[14-16] Despite these advantages of FIT over gFOBT, gFOBT is still being used in several regions.[5] Although interval cancer rate is a key quality indicator in screening programs, data on the incidence rate of FOBT iCRC is limited and no data are available on how these two types of FOBT compare with regard to iCRC.

We therefore performed a systematic literature search and meta-analysis to determine and compare the pooled incidence rates of iCRC following gFOBT and FIT in population-based CRC screening programs. Secondly, we assessed if screening-related or patient-related factors are associated with iCRC incidence rate.

METHODS

Search strategy and selection criteria

We carried out a systematic review and meta-analysis of published trials, including observational and experimental trials or both, according to the preferred reporting items

for systematic reviews and meta-analyses (PRISMA) guidelines.[17] We additionally used a checklist containing specifications for reporting of meta-analysis of observational studies in epidemiology (MOOSE).[18]

Studies were identified through a systematic literature search until the 10th of May 2016 in the following electronic databases: Medline, Embase, Web of Science, the Cochrane Library, PubMed publisher and Google Scholar. A search update was performed on the 12th of December 2017. The search strategy was designed and conducted using controlled vocabulary supplemented with key words and without any restrictions on date or language (supplementary material 1). The titles and abstracts of identified studies were reviewed by at least two of the authors independently (EW, EHS or EJG). Studies were excluded that did not address the research question, based on the inclusion and exclusion criteria mentioned below. The full texts of the remaining publications were carefully and independently examined by the same authors. In case of disagreement, consensus was reached by consulting a fourth author (MCWS). In addition, the reference lists of the included studies were hand-searched to identify additional, potentially relevant studies (published within 5 years preceding our search).

Studies were included if they reported on CRC occurrence within one to five years after a negative gFOBT or FIT in average-risk screening populations. Both prospective and retrospective studies were included. Only studies comprising asymptomatic average-risk individuals aged 40 years and above were included, as these were considered representative for a population-based CRC screening program. Studies were eligible if participants with a positive test were referred for endoscopic confirmation. For the purpose of this systematic review, diagnostic tests accepted as endoscopic confirmation included colonoscopy, or if colonoscopy was not available or contra/-indicated sigmoidoscopy, computed tomography colonoscopy or double contrast barium enema. Only full-text articles were included. We did not restrict studies based on language or publication date. If the same screening cohort was described in more than one publication, the one with the most recently updated and most complete data was included.

Accuracy studies in which all participants underwent both the stool test and colonoscopy were excluded. Also excluded were reviews, systematic reviews, editorials and letters to the editor. Lastly, studies in which individuals were referred to endoscopy after two or more consecutive positive tests were excluded.

Outcomes

The primary outcome was the pooled incidence of interval colorectal cancer during gFOBT and FIT screening per 100,000 person-years (p-y) in an average CRC screening population.

Secondary outcomes were the proportional rate between iCRCs en screen-detected CRCs and pooled incidence of iCRCs per subgroup. Subgroups were categorized by means of screening-related and screenee-related factors, including number of screening round, duration of follow-up after a negative stool test, positivity cut-off, gender, age, tumour stage, and tumour location.

Definitions

Screen-detected CRC was defined as a CRC detected by endoscopic conformation after a positive test. Interval colorectal cancers were defined in agreement with the definition of the World Endoscopy Organization as cancers diagnosed after a negative test and before the next test was due.[7] Post-colonoscopy CRCs diagnosed after a negative colonoscopy were not taken into account. If a study did not describe when the next occult blood test was due, we assumed the interval to be 2 years. Proximal CRCs were defined as CRCs located in the cecum, ascending colon, transverse colon, or splenic flexure; and distal CRCs as CRCs located in the descending colon, sigmoid colon, or rectum. Early CRC was defined as Dukes A, or TNM stage 1.[19] In case of quantitative FIT, we converted units for positivity cut-off into micrograms (µg) of haemoglobin (Hb) per gram of stool for each study.[20]

Data extraction

Study characteristics and data were independently extracted by two investigators (EW and EHS) and recorded on a standardized data extraction form. Any discrepancies were resolved by consensus. The types of data extracted are shown in supplementary material 2. If data were incomplete, the corresponding author was asked to provide the missing information. Alternatively, we derived data from other publications on the same study cohort. If applicable, data from multiple screening rounds were included for analysis.

Data analyses

Incidence rates of iCRC were calculated per 100,000 p-y. The follow-up p-y were calculated as the number of participants with a negative test multiplied by the mean years of follow-up or the number of years for which interval cancers were identified, by using data from the cancer registry.

Pooled incidence rates were obtained by fitting random effect Poisson regression models. Heterogeneity was quantified by using the inconsistency index (I^2) test. Heterogeneity values ranged from 0% (no heterogeneity) to 100%.[21, 22] I^2 greater than 25%, 50%, and 75% was defined as indicative of low, moderate, and high heterogeneity.[22] For studies describing both a gFOBT and FIT study arm, we interpreted each arm as a separate study, ignoring the within study correlation. We used prediction intervals to calculate the expected incidence of iCRC for new settings similar to the ones included in the meta-analysis.

An incidence rate ratio was used to compare pooled incidence rates of iCRC after FIT and gFOBT. The sensitivity analysis included only studies in which both a FIT and gFOBT study arm was described. This allowed comparison of incidence rate of iCRC between the two test types in the same study population.

Meta-regression analysis

Meta-regression analyses served to assess if screening-related and patient-related factors were associated with FOBT iCRCs. For these analyses, gFOBT and FIT studies were pooled together as only few studies reported on these factors. Incidence rate ratios or proportions were used to describe categorical variables. Relative risk was used for continuous outcomes.

Quality assessment

A funnel plot was created to assess the presence of publication bias.[23] The study quality of observational studies was assessed using the Ottawa Newcastle criteria of Wells *et al*.[24] Studies were considered as high quality studies in case of a score of eight or nine out of nine stars according to the Ottowa Newcastle criteria, absence of selection bias and adequate cohort follow-up. Selection bias was considered to be present if <90% of the total inception cohort was followed. With respect to study follow-up a minimum of 2 years follow-up was required to define a high-quality study. The study quality of randomised trials was assessed using the Cochrane risk of bias tool.[25] We performed a post hoc subset analysis with high-quality studies only, to assess incidence of iCRC. The quality of evidence was rated by the Grading of Recommendations Assessment, Development and Evaluation (GRADE).[26]

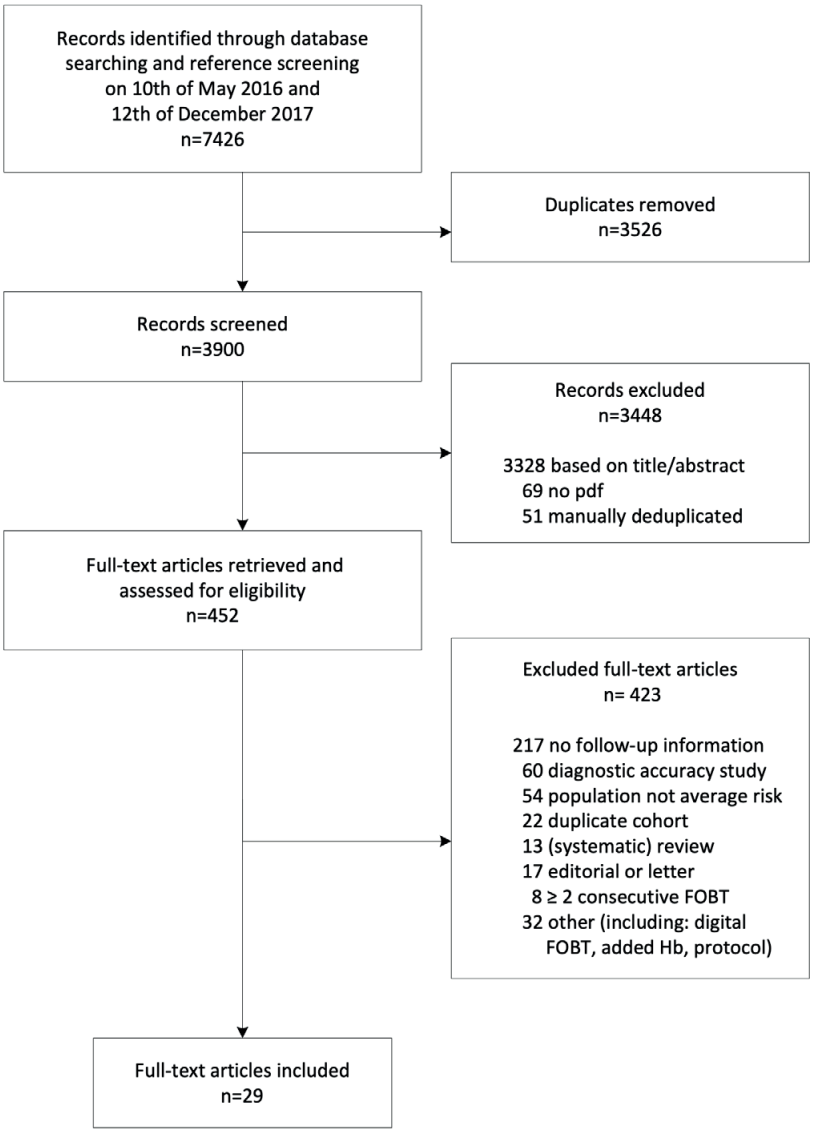
All analyses were done using R version 2.15.1.

RESULTS

In total, 7426 records were identified. After removal of duplicates, 3526 records were screened for eligibility based on title and/or abstract. In total, 452 full-text records were reviewed for eligibility criteria, of which 423 were excluded for various reasons (Figure 1). Thus, twenty-nine studies were included for qualitative and quantitative analysis.[2, 3, 6, 27-52]

Characteristics of the included studies are shown in Table 1. Nineteen studies were performed in Europe, seven in Asia, and three in North America. Fourteen studies contained data on FIT related iCRCs, twelve on gFOBT related iCRCs, and three on both gFOBT and FIT related iCRCs. The median faecal haemoglobin (Hb) positivity cut-off in the 17 studies that provided FIT results, was 20 (range 10-200) microgram Hb/gram faeces. The study quality score of the twenty-seven observational studies ranged from 4 to 8 stars according to the

Figure 1 Flow chart of literature search and study inclusion



FOBT: faecal occult blood test; Hb: haemoglobin

Ottawa Newcastle criteria (Supplementary Table 1). The two randomised controlled trials were both scored as good quality studies according to the Cochrane risk of bias tool.[32, 39]

Meta-analysis

Meta-analysis comprised data of 6,987,825 screening participants with a negative FOBT result, ranging from 1071 to 2,033,526 participants per study (Table 1). Of all twenty-nine

Table 1 Study and test characteristics of a) FIT studies and b) gFOBT studies included in meta-analysis.

a) FIT studies

Study	Country	Time period	Age range of population screened	Males in population screened	No. of screening rounds included in meta-analyses	No. of stools/ no. of samples per stool	FIT cut-off	Participants with a negative FIT	Person years	Total screen-detected CRCs	Total FIT iCRCs
			years	%			µg Hb/g faeces	n		n	n
Chen[28]	Taiwan	1994-2008	>40	n.a.	1	n.d./1	20	221,874	443,748	298	133
Chiu[29]	Taiwan	2004-2008	50-70	38	3	1/1	20	1,113,932	6,683,592	2728	968
Crotta[30]	Italy	2001-2008	50-74	n.a.	4	1/1	20	1928	16,388	8	5
Denters*[32]	NL	2006-2008	50-74	45	1	1/1	10	2638	5276	21	4
Digby[33]	Scotland	2010-2011	50-74	n.a.	1	n.d./1	80	30,140	60,280	30	31
Giorgi Rossi[50]	Italy	2000-2008	50-79	46	1	1/1	20	805,914	805,914	n.a.	172
Itoh[35]	Japan	1991-1992	≥40	n.a.	1	1/1	10	26,370	52,740	77	12
Jensen[36]	USA	2007-2008	50-69	47	3	n.d./1	20	641,559	2,566,236	830	242
Launoy[38]	France	2001-2003	50-74	43	1	2/1	≥67 in ≥1 FITs	6987	13,974	24	4
Levi*[39]	Israel	2008-2011**	50-75	45	1	3/1[40]	≥14 in ≥1 FITs	1071	2142	6	0
McNamara[41]	Ireland	2008-2012	50-75	42	1	2/1	≥20 in ≥1 FITs	4549	9098	21	1
Nakama[42]	Japan	1991	>40	49	1	1/1	Qualitative	3208	9624	10	4
Parente[44]	Italy	2005-2007	50-69	n.a.	1	1/1	100	36,401	72,802	165	8
Portillo[51]	Spain	2009-2015	50-69	n.a.	3	1/1	17-20	769,124	4,614,744	2518	186
Shin[46]	Korea	2004-2007	>50	54	2	1/1	10 or qualtitative	2,033,526	8,134,104	2961	2047
van der Vlugt[52]	NL	2006-2015	50-74	n.a.	3	1/1	10	15,711	111,705	89	27
Zappa*[48]	Italy	1992-1997	50-70	n.a.	2	n.d./1	200-300	20,120	80,480	73	8
Total	-	-	-	-	-	-	-	**5,725,052**	**23,682,847**	**9859**	**3852**

b) gFOBT studies

Study	Country	Time period	Age range of population screened	Males in population screened	No. of screening rounds included in meta-analyses	N. of stool / n. of samples per stool	Participants with a negative gFOBT	Person years	Total screen-detected CRCs	Total gFOBT iCRCs
			years	%			n		n	n
Blom[49]	Sweden	2008-2014	60-69	n.a.	3	3/1	193,690	1,162,140	219	301
Bouvier[27]	France	1991-1994	45-74	n.a.	1	n.d.	69,287	207,861	152	100
Cummings[31]	USA	1984	>40	40	1	3/2	11,233	22,466	13	1
Denters*[32]	NL	2006-2008	50-74	42	1	3/2	2059	4118	8	4
Faivre[34]	France	1988-1998	45-74	n.a.	6	3/2	131,680	263,360	196	219
Hardcastle[2]	UK	1981-1991	50-74	48	6	3/1 or 2	43,748	341,234	236	147
Kewenter[37]	Sweden	1982-1985	60-64	n.a.	1	3/2	8700	14,503	35	16
Kronborg[3]	Denmark	1985-1995	45-75	47	5	3/2	85,794	171,588	120	146
Levi*[39]	Israel	2008-2011**	50-75	43	1	3/2	2178	4356	8	5
Mandel[40]	USA	1957-1982	50-80	n.a.	5	3/2	91,332	456,660	183	22
Paimela[43]	Finland	2004-2006	60-64	31	1	3/2	36,708	70,357	95	35
Rennert[45]	Israel	1992	50-74	n.a.	1	3/2	21,158	60,124	58	10
Souques[47]	France	1980-1995	40-70	n.a.	7	3/2	24,504	171,528	15	10
Steele[6]	Scotland	2000-2007	50-69	45	3	3/2	498,724	2,992,344	698	635
Zappa*[48]	Italy	1992-1997	50-70	n.a.	2	n.d.	31,978	127,912	66	45
Total	-	-	-	-	-	-	**1,252,773**	**6,070,551**	**2073**	**1696**

*Studies that described both a FIT and gFOBT study arm

**Exact study time period not described. However, this study was approved in 2008 and published in 2011.

n.a.: not applicable; NL: Netherlands; USA: United states of America; FOBT: faecal occult blood test; Hb: haemoglobin; CRC: colorectal cancer; iCRC: interval colorectal cancer; n.d.: not described; FIT: faecal immunochemical test; gFOBT guaiac faecal occult blood test

included studies, total follow-up for participants with a negative screening test was 32 million p-y, with a mean follow-up of 4.0 years. In these studies, 11,932 screen-detected CRCs (range 6 to 2961) and 5548 iCRCs (range 0 to 2047) were documented. For every iCRC, 2.6 screen-detected CRC were found with FIT. In gFOBT-based studies the ratio between iCRC and screen-detected CRC was 1:1.2. The Forest plot of the ratio of iCRC following a negative stool test compared to screen-detected CRC is shown in Supplementary Figure 1. This ratio was 0.19 (95%CI 0.13 to 0.27) for FIT studies compared to 0.36 (95%CI 0.28 to 0.45) for gFOBT studies, p=0.005, I^2=99%.

The overall pooled incidence rate of iCRC following a negative stool test was 26 (95%CI 19 to 36; I^2=99%, n=29 studies) per 100,000 p-y (Figure 2). Pooled incidence rates of iCRC for FIT and gFOBT were 20 (95%CI 14 to 29; I^2=99%) and 34 (95%CI 20 to 57; I^2=99%) per 100,000 p-y, respectively (Figure 2). The pooled incidence rate ratio between FIT iCRC and gFOBT iCRC was 0.58 (95%CI 0.32-1.07, n=29 studies). The GRADE level of evidence was very low (Supplementary Table 2). The funnel plot provided no evidence for the presence of publication bias, (Supplementary Figure 2). The prediction intervals of the incidence rate of FIT and gFOBT iCRC are shown in Figure 2.

Subgroup analysis of the studies with high quality established with the Ottawa Newcastle criteria and Cochrane risk of bias tool yielded an incidence rate of iCRC after FIT of 15 (95%CI 8 to 30, n=7 studies) and after gFOBT of 55 (95%CI 35 to 87, n=8 studies) per 100,000 p-y.

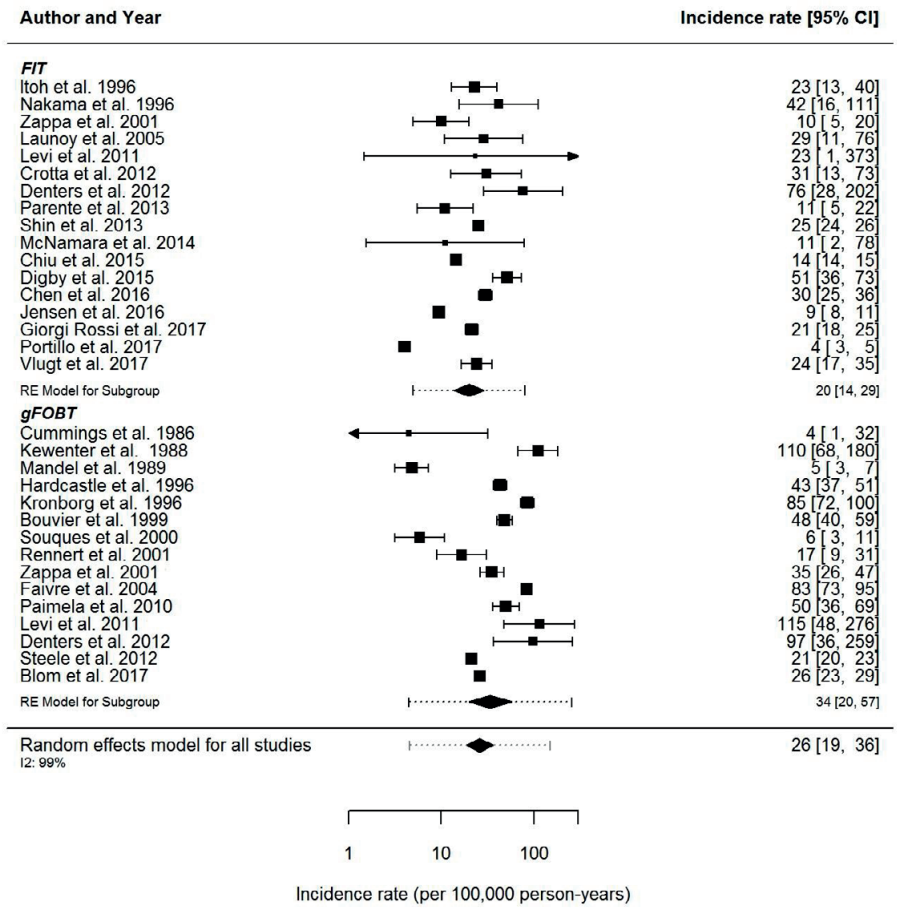
Three studies that described both a FIT and gFOBT arm were included in a sensitivity analysis to compare incidence rate of iCRC between FIT and gFOBT.[32, 39, 48] The pooled incidence rate ratio between FIT iCRC and gFOBT iCRC was 0.36 (95% CI 0.17-0.75, I^2=10%). This ratio was classified as high-quality evidence according to the GRADE score (Supplementary Table 2).

Meta-regression analyses

For meta-regression analyses, data of FIT and gFOBT studies were pooled.

Fifteen out of twenty-nine studies provided data on test related iCRCs after the first screening round (Table 2). Five studies also provided data on iCRC after the second round, four studies after the third, and one study after four rounds of screening. After the third round there was a significant lower risk of iCRC after a negative test compared to the first round (Table 2).

Figure 2 Forest plot showing the incidence rate of interval colorectal cancers per study and summary estimates for FIT and gFOBT



|·····| Prediction interval

gFOBT: guaiac faecal occult blood test; FIT: faecal immunochemical test; CI: confidence interval

Eight out of twenty-nine studies provided data on iCRCs one year after a negative test. Six studies provided data on iCRC two years after a negative test and two studies three years after a negative test (Table 2). Compared to one year after a negative FOBT, the relative risk of iCRC was 1.25 (95%CI 1.05-1.49) after two and 1.19 (95%CI 0.89-1.13) after three years.

Thirteen out of seventeen FIT studies used a single quantitative positivity cut-off, ranging from 0-100 µg Hb/g faeces. Association between FIT cut-off and FIT iCRC yielded a relative risk of developing FIT iCRC of 1.00 (0.89-1.13) per 10 µg Hb/g faeces cut-off increase (Table 2).

Table 2 Relative risk to develop FOBT iCRC per screening round, in years since last negative test, and cut-off

Subgroup variable	Studies n	Relative risk (95%CI)	Study references
Screening round			
1	15	*Reference*	6, 27, 28, 31, 32, 36-39, 42, 44-46, 51, 52
2	5	0.93 (0.85-1.02)	6, 36, 45, 46, 52
3	4	0.76 (0.66-0.88)	6, 36, 45, 52
4	1	0.77 (0.54-1.10)	36
Time in years since last negative test			
1	8	*Reference*	2, 27, 28, 38, 42, 45, 50
2	6	1.25 (1.05-1.57)	2, 27, 28, 38, 42, 50
3	2	1.19 (0.76-1.87)	27, 42
Cut-off*			
Per 10 µg Hb/g faeces increase	17	1.00 (0.89-1.13)	28-30, 32, 33, 35, 36, 38, 39, 41, 44, 50, 52

*Based on thirteen studies that used a single fixed cut-off.

Table 3 Incidence rate ratios of FOBT iCRC per gender and age

Subgroup	Studies n	Incidence rate ratio (95%CI)	I^{2}*	Study references
Gender**	3	1.22 (0.94-1.57)	0%	33, 43, 50
Age***	2	0.25 (0.09-0.65)	62%	33, 50

*Heterogeneity was quantified by the heterogeneity variance, using the inconsistency index (I^2) test (range from 0% to 100%). We regarded values greater than 25%, 50%, and 75% for the I^2 as indicative of low, moderate, and high statistical heterogeneity, respectively.
**male versus female
***<60 years versus ≥60

Based on three studies, the iCRC incidence rate ratio between males and females was 1.22 (95%CI 0.94-1.57, I^2= 0%). And based on two out of twenty-nine studies, the iCRC incidence rate ratio of screenees <60 years of age to screenees ≥60 years was 0.25 (95%CI 0.09-0.65, I^2=62%) (Table 3).

Eight out of twenty-nine studies described the location of iCRCs in the colon.[6, 27, 31, 33, 39, 50-52] Based on these studies, iCRCs were located distal from the splenic flexure in 67% (95%CI 64%-70%, I^2=0%) of cases. Six out of twenty-nine studies described tumour stages of iCRCs.[6, 27, 31, 33, 51, 52] These iCRCs were staged as early CRCs in 22% (95%CI 17%-28%, I^2=62%) of cases.

DISCUSSION

This is the first systematic review and meta-analysis to estimate the pooled incidence rates of interval colorectal cancers following negative FOBTs in a CRC screening setting. It showed that iCRCs occur in both FIT-based and gFOBT-based CRC screening. However, the incidence of iCRC is higher after a negative gFOBT than after a negative FIT.

Interval cancer rates reflect the sensitivity of a screening test and quality of a screening program. International guidelines, therefore, designate the interval cancer rate as an important outcome measure.[8, 14] Pooled data on incidence of interval cancers following a negative gFOBT and FIT have been awaited for some time.[14, 53]

The findings of this meta-analysis emphasize that screenees should be adequately informed about the risk of CRC after a negative occult blood test. They may mistakenly feel disease-free and fail to respond to CRC symptoms.[54] A Swedish study indeed reported a significant delay in CRC diagnosis among those with a false-negative FOBT.[55] However, recent evidence showed that both overall and CRC-specific survival rates were better for gFOBT interval cancers than for cancers arising in a non-screened population.[6] We found that iCRC accounted for a significant proportion of CRC found in both gFOBT-based and FIT-based screening programs. In the included gFOBT studies, the total number of CRCs missed by gFOBT almost equalled the number of screen-detected CRCs.

The higher incidence of iCRCs after a negative gFOBT compared to FIT in sensitivity analysis is likely due to the higher sensitivity of FIT for the detection of haemoglobin. Best evidence suggests that the most used gFOBT probably has an effective cut-off of around 150 µg of Hb/g of faeces, whereas the accurate detection level of most FITs lies at 5-10 µg Hb/g of faeces.[13, 56, 57] The incidence of iCRC as primary outcome did not significantly differ between gFOBT and FIT. This may have been due to excessive study bias as shown by subgroup analyses.

We found that older age was associated with a higher iCRC incidence after a negative test. Indeed, the elderly have a higher risk of CRC and its precursors.[58-60] Further, the risk of a FOBT-related iCRC was not significantly different between males versus females. This implies, in view of the fact that FOBT-screening detects more CRCs in males[61], that the ratio of screen-detected colorectal cancers versus interval cancers is less favourable in women than in men. Furthermore, we found lower risks of FOBT-related iCRC with every screening round compared to the first round. A possible explanation for this finding is that with every screening round more CRCs are detected and therefore changes of missing CRC with FIT decline as well. Lastly, use of a higher positivity cut-off resulted in a similar incidence of FIT interval cancers. This finding needs to be confirmed when more data become available.

Moreover, studies included in our analyses used low cut-offs which might be the reason that this association was not found.

For quality assessment of CRC screening, it is recommended to monitor the iCRC incidence.[53] Various measures can be used for this purpose. The incidence of FOBT iCRC can be calculated as the ratio of iCRCs versus i) participants with a negative test; ii) person-years follow-up in those with a negative test; iii) total CRCs (detected and missed); and (iv) CRCs expected without screening. In our study we assessed iCRC per person-years, reflecting the absolute number of iCRC cases in CRC screening populations over time. However, when calculating incidence rates on a program level, participation rates should also be taken into account. Secondary outcome in our study was the relative rate of iCRC versus screen-detected CRC, which is an indirect measure of test sensitivity. Previously published data revealed a higher test sensitivity for FIT compared to gFOBT.[62, 63]

Although this comprehensive meta-analysis is based on a large number of person-years, the point estimates of our calculated pooled incidence rates of iCRC should be interpreted cautiously. First, high statistical heterogeneity among studies was shown. We assessed the robustness of conclusions concerning the effect sizes of real interest in our meta-analysis as substantial statistical heterogeneity was observed in the overall pooled data. Statistical heterogeneity represents the approximate proportion of total variability in point estimates that can be attributed to heterogeneity in underlying incidence rates. To explain the observed heterogeneity of the incidence rates between studies we performed subgroup analyses. A potential important source of heterogeneity are differences between populations screened in terms of gender distribution, age distribution, and number of performed screening rounds. These were all identified in the subgroup analyses as factors that partly explained the observed heterogeneity. Furthermore, we performed sensitivity analyses by only including studies directly comparing gFOBT and FIT, limiting the influence of factors introducing heterogeneity, to directly estimate the difference in incidence rates of both tests. This analysis resulted in a higher gFOBT iCRC incidence compared to FIT. The marked inconsistency among the included trials in incidence rate ratios for iCRC (I^2 = 99%) was substantially reduced (I^2 = 10%) when differences between populations were taken into account. Another important factor potentially introducing heterogeneity between studies was study quality. Additional sensitivity analyses of high quality studies only showed a significant higher gFOBT iCRC incidence than FIT iCRC incidence. Second, test kits may be discrepant in terms of cancer detectability and the resultant future risk of iCRC. Test reliability, stability and the ability to detect invasive cancer or advanced adenoma of different kits have been compared in several previous studies.[64, 65] Further stratifying analysis in our study to correct for differences in test kits was not feasible. Only few studies

have reported on iCRCs within specific subgroups, therefore limited analyses could be done for gFOBT and FIT separately.

In conclusion, interval cancers occur in both gFOBT and FIT-based CRC screening programs. The latter is associated with a significantly lower incidence of iCRC, which further supports the use of FIT over gFOBT.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359-86.
2. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-77.
3. Kronborg O, Fenger C, Olsen J, Jørgensen OD. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348(9040):1467-71.
4. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, 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):1365-71.
5. 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.
6. Steele RJ, McClements P, Watling C, Libby G, Weller D, Brewster DH, et al. Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. Gut. 2012;61(4):576-81.
7. Sanduleanu S, le Clercq CM, Dekker E, Meijer GA, Rabeneck L, Rutter MD, et al. Definition and taxonomy of interval colorectal cancers: a proposal for standardising nomenclature. Gut. 2015;64(8):1257-67.
8. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation of screening outcomes. Endoscopy. 2012;44 Suppl 3:SE49-64.
9. Dekker E, Sanduleanu S. Colorectal cancer: Strategies to minimize interval CRC in screening programmes. Nat Rev Gastroenterol Hepatol. 2016;13(1):10-2.
10. Singh S, Singh PP, Murad MH, Singh H, Samadder NJ. Prevalence, risk factors, and outcomes of interval colorectal cancers: a systematic review and meta-analysis. Am J Gastroenterol. 2014;109(9):1375-89.
11. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362(19):1795-803.
12. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100(7):1103-10.
13. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327-37.
14. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
15. Kuipers EJ, Spaander MC. Colorectal Cancer Screening by Colonoscopy, CT-Colonography, or Fecal Immunochemical Test. J Natl Cancer Inst. 2016;108(2):djv383.
16. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol. 2013;10(3):130-42.
17. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9, W64.

18. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000;283(15):2008-12.
19. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol. 2010;17(6):1471-74.
20. Fraser CG, Allison JE, Halloran SP, Young GP, Org WE. A Proposal to Standardize Reporting Units for Fecal Immunochemical Tests for Hemoglobin. Jnci-J Natl Cancer I. 2012;104(11):810-14.
21. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88.
22. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
23. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. BMJ. 2001;323(7304):101-5.
24. Wells G, Shea B, O'connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm; 2010.
25. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
26. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94.
27. Bouvier V, Launoy G, Herbert C, Lefevre H, Maurel J, Gignoux M. Colorectal cancer after a negative Haemoccult II test and programme sensitivity after a first round of screening: the experience of the Department of Calvados (France). Br J Cancer. 1999;81(2):305-9.
28. Chen CH, Tsai MK, Wen CP. Extending Colorectal Cancer Screening to Persons Aged 40 to 49 Years With Immunochemical Fecal Occult Blood Test: A Prospective Cohort Study of 513,283 Individuals. J Clin Gastroenterol. 2016;50(9):761-68.
29. Chiu HM, Chen SL, Yen AM, Chiu SY, Fann JC, Lee YC, et al. Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. Cancer. 2015;121(18):3221-9.
30. Crotta S, Segnan N, Paganin S, Dagnes B, Rosset R, Senore C. High rate of advanced adenoma detection in 4 rounds of colorectal cancer screening with the fecal immunochemical test. Clin Gastroenterol Hepatol. 2012;10(6):633-8.
31. Cummings KM, Michalek A, Tidings J, Herrera L, Mettlin C. Results of a public screening program for colorectal cancer. N Y State J Med. 1986;86(2):68-72.
32. Denters MJ, Deutekom M, Bossuyt PM, Stroobants AK, Fockens P, Dekker E. Lower risk of advanced neoplasia among patients with a previous negative result from a fecal test for colorectal cancer. Gastroenterology. 2012;142(3):497-504.
33. Digby J, Fraser CG, Carey FA, Lang J, Stanners G, Steele RJ. Interval cancers using a quantitative faecal immunochemical test (FIT) for haemoglobin when colonoscopy capacity is limited. J Med Screen. 2015;23(3):130-34.
34. Faivre J, Dancourt V, Lejeune C, Tazi MA, Lamour J. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology. 2004;126(7):1674-80.
35. Itoh M, Takahashi K, Nishida H, Sakagami K, Okubo T. Estimation of the optimal cut off point in a new immunological faecal occult blood test in a corporate colorectal cancer screening programme. J Med Screen. 1996;3(2):66-71.

36. Jensen CD, Corley DA, Quinn VP, Doubeni CA, Zauber AG, Lee JK, et al. Fecal immunochemical test program performance over 4 rounds of annual screening: A retrospective cohort study. Ann Intern Med. 2016;164(7):456-63.
37. Kewenter J, Bjork S, Haglind E, Smith L, Svanvik J, Ahren C. Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27,700 subjects. Cancer. 1988;62(3):645-51.
38. Launoy GD, Bertrand HJ, Berchi C, Talbourdet VY, Guizard AV, Bouvier VM, et al. Evaluation of an immunochemical fecal occult blood test with automated reading in screening for colorectal cancer in a general average-risk population. Int J Cancer. 2005;115(3):493-6.
39. Levi Z, Birkenfeld S, Vilkin A, Bar-Chana M, Lifshitz I, Chared M, et al. A higher detection rate for colorectal cancer and advanced adenomatous polyp for screening with immunochemical fecal occult blood test than guaiac fecal occult blood test, despite lower compliance rate. A prospective, controlled, feasibility study. Int J Cancer. 2011;128(10):2415-24.
40. Mandel, Bond, Bradley ea. Sensitivity Specificity and Positive Predictivity of the Hemoccult Test in Screening for Colorectal Cancers the University of Minnesota's Colon Cancer Control Study USA. Gastroenterology1989. p. 597-600.
41. McNamara D, Leen R, Seng-Lee C, Shearer N, Crotty P, Neary P, et al. Sustained participation, colonoscopy uptake and adenoma detection rates over two rounds of the Tallaght-Trinity College colorectal cancer screening programme with the faecal immunological test. Eur J Gastroenterol Hepatol. 2014;26(12):1415-21.
42. Nakama H, Kamijo N, Abdul Fattah AS, Zhang B. Validity of immunological faecal occult blood screening for colorectal cancer: a follow up study. J Med Screen. 1996;3(2):63-5.
43. Paimela H, Malila N, Palva T, Hakulinen T, Vertio H, Jarvinen H. Early detection of colorectal cancer with faecal occult blood test screening. Br J Surg. 2010;97(10):1567-71.
44. Parente F, Boemo C, Ardizzoia A, Costa M, Carzaniga P, Ilardo A, et al. Outcomes and cost evaluation of the first two rounds of a colorectal cancer screening program based on immunochemical fecal occult blood test in northern Italy. Endoscopy. 2013;45(1):27-34.
45. Rennert G, Rennert HS, Miron E, Peterburg Y. Population colorectal cancer screening with fecal occult blood test. Cancer Epidemiol Biomarkers Prev. 2001;10(11):1165-8.
46. Shin A, Choi KS, Jun JK, Noh DK, Suh M, Jung KW, et al. Validity of fecal occult blood test in the national cancer screening program, Korea. PLoS One. 2013;8(11):e79292.
47. Souques M, Zummer K. The Hemoccult II test: results of 16 years of screening tests at the Tumor Prevention Service of the City of Paris. Presse Med. 2000;29(18):983-6.
48. Zappa M, Castiglione G, Paci E, Grazzini G, Rubeca T, Turco P, et al. Measuring interval cancers in population-based screening using different assays of fecal occult blood testing: The district of florence experience. Int J Cancer. 2001;92(1):151-4.
49. Blom J, Tornberg S. Interval cancers in a guaiac-based colorectal cancer screening programme: Consequences on sensitivity. J Med Screen. 2017;24(3):146-52.
50. Giorgi Rossi P, Carretta E, Mangone L, Baracco S, Serraino D, Zorzi M, et al. Incidence of interval cancers in faecal immunochemical test colorectal screening programmes in Italy. J Med Screen. 2017:969141316686391.
51. Portillo I, Arana-Arri E, Idigoras I, Bilbao I, Martínez-Indart L, Bujanda L, et al. Colorectal and interval cancers of the Colorectal Cancer Screening Program in the Basque Country (Spain). World J Gastroenterol. 2017;23(15):2731-42.
52. van der Vlugt M, Grobbee EJ, Bossuyt PMM, Bos A, Bongers E, Spijker W, et al. Interval Colorectal Cancer Incidence Among Subjects Undergoing Multiple Rounds of Fecal Immunochemical Testing. Gastroenterology. 2017;153(2):439-47.e2.

53. Wieten E, Spaander MC, Kuipers EJ. Accrediting for screening-related colonoscopy services: What is required of the endoscopist and of the endoscopy service? Best Pract Res Clin Gastroenterol. 2016;30(3):487-95.
54. van Dam L, Bretthauer M. Ethical issues in colorectal cancer screening. Best Pract Res Clin Gastroenterol. 2014;28(2):315-26.
55. Gillberg A, Ericsson E, Granstrom F, Olsson LI. A population-based audit of the clinical use of faecal occult blood testing in primary care for colorectal cancer. Colorectal Dis. 2012;14(9):e539-46.
56. Allison JE, Sakoda LC, Levin TR, Tucker JP, Tekawa IS, Cuff T, et al. Screening for colorectal neoplasms with new fecal occult blood tests: update on performance characteristics. J Natl Cancer Inst. 2007;99(19):1462-70.
57. Moss S, Mathews C, Day TJ, Smith S, Seaman HE, Snowball J, et al. Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England. Gut. 2017;66(9):1631-44.
58. 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):1585-89.
59. Ferlitsch M, Reinhart K, Pramhas S, Wiener C, Gal O, Bannert C, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA. 2011;306(12):1352-58.
60. Wieten E, Schreuders EH, Nieuwenburg SA, Hansen BE, Lansdorp-Vogelaar I, Kuipers EJ, et al. Effects of Increasing Screening Age and Fecal Hemoglobin Cutoff Concentrations in a Colorectal Cancer Screening Program. Clin Gastroenterol Hepatol. 2016;14(12):1771-77.
61. McDonald PJ, Strachan JA, Digby J, Steele RJ, Fraser CG. Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. Clin Chem Lab Med. 2012;50(5):935-40.
62. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2008;149(9):638-58.
63. Zhu MM, Xu XT, Nie F, Tong JL, Xiao SD, Ran ZH. Comparison of immunochemical and guaiac-based fecal occult blood test in screening and surveillance for advanced colorectal neoplasms: a meta-analysis. J Dig Dis. 2010;11(3):148-60.
64. De Girolamo G, Goldoni CA, Corradini R, Giuliani O, Falcini F, Sassoli De'Bianchi P, et al. Ambient temperature and FIT performance in the Emilia-Romagna colorectal cancer screening programme. J Med Screen. 2016;23(4):186-91.
65. 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.

Supplementary Material 1 Search strategy

Embase

('large intestine tumor'/exp OR ((colorect* or colon* or rect* or anal* or anus* or intestin* or bowel*) NEAR/3 (carcinom* or neoplas* or adenocarcinom* or cancer* or tumor* or tumour* or sarcom* or polyp* or adenom*)):de,ab,ti) AND ((((('occult blood'/exp OR 'occult blood':de,ab,ti) AND (faecal or fecal or feces or faeces or stool*):de,ab,ti) OR (FOBT* or FIT* or gFOBT*):de,ab,ti) AND ((immunohistochem* or immunochem* or immunol* or guaiac*):de,ab,ti OR immunochemistry/exp OR guaiac/exp) OR (('fecal immunochemical' NEXT/1 test* or 'faecal immunochemical' NEXT/1 test* or 'fecal immunochemistry' NEXT/1 test* or 'faecal immunochemistry' NEXT/1 test* or ColoScreen or Hema-Screen or Hemdetect or Hemoccult or SENSA or Hema-Check or HemaCheck or hemoCARE or Peroheme or ColoCare or Lifeguard or Fecatwin or HemaWipe or Instaccult or Monohaem or Okokit or Seracult or Dencoccult or Early-detector or Earlydetector or Fe-Cult or Fecult or Feca-EIA or FecaEIA or Hemo-FEC or HemoFEC or Hexagon or SureScreen or Hemaprompt or Hemdetect or Camco-PAK or CamcoPAK or Colocheck or Cecogenics or Hematest or Dencocult or Fecatest or Hemofecia or Quick-CULT or QuickCULT)):de,ab,ti)

Medline Ovid

(exp Colorectal Neoplasms/ OR ((colorect* or colon* or rect* or anal* or anus* or intestin* or bowel*) adj3 (carcinom* or neoplas* or adenocarcinom* or cancer* or tumor* or tumour* or sarcom* or polyp* or adenom*)).mp.) AND ((((exp Occult Blood/ OR occult blood.mp.) AND (faecal or fecal or feces or faeces or stool*).mp.) OR (FOBT* or FIT* or gFOBT*).mp.) AND ((immunohistochem* or immunochem* or immunol* or guaiac*).mp. OR exp Immunochemistry/ OR exp Guaiac/) OR ((fecal immunochemical test* or faecal immunochemical test* or fecal immunochemistry test* or faecal immunochemistry test* or ColoScreen or Hema-Screen or HemaScreen or Hemdetect or Hemoccult or SENSA or Hema-Check or HemaCheck or hemoCARE or Peroheme or ColoCare or Lifeguard or Fecatwin or HemaWipe or Instaccult or Monohaem or Okokit or Seracult or Dencoccult or Early detector or Earlydetector or Fe Cult or Fecult or Feca EIA or FecaEIA or Hemo FEC or HemoFEC or Hexagon or SureScreen or Hemaprompt or Hemdetect or Camco PAK or CamcoPAK or Colocheck or Cecogenics or Hematest or Dencocult or Fecatest or Hemofecia or Quick-CULT or QuickCULT)).mp.)

Cochrane Library

(((colorect* or colon* or rect* or anal* or anus* or intestin* or bowel*) NEAR/3 (carcinom* or neoplas* or adenocarcinom* or cancer* or tumor* or tumour* or sarcom* or polyp* or adenom*)):kw,ab,ti) AND ((((('occult blood':kw,ab,ti) AND (faecal or fecal or feces or faeces or stool*):kw,ab,ti) OR (FOBT* or FIT* or gFOBT*):kw,ab,ti) AND ((immunohistochem* or immunochem* or immunol* or guaiac*):kw,ab,ti) OR (('fecal immunochemical' NEAR/1

test* or 'faecal immunochemical' NEAR/1 test* or 'fecal immunochemistry' NEAR/1 test* or 'faecal immunochemistry' NEAR/1 test* or ColoScreen or Hema-Screen or Hemdetect or Hemoccult or SENSA or Hema-Check or HemaCheck or hemoCARE or Peroheme or ColoCare or Lifeguard or Fecatwin or HemaWipe or Instaccult or Monohaem or Okokit or Seracult or Dencoccult or Early-detector or Earlydetector or Fe-Cult or Fecult or Feca-EIA or FecaEIA or Hemo-FEC or HemoFEC or Hexagon or SureScreen or Hemaprompt or Hemdetect or Camco-PAK or CamcoPAK or Colocheck or Cecogenics or Hematest or Dencocult or Fecatest or Hemofecia or Quick-CULT or QuickCULT)):kw,ab,ti)

Science Citation Index

TS=(((((colorect* or colon* or rect* or anal* or anus* or intestin* or bowel*) NEAR/3 (carcinom* or neoplas* or adenocarcinom* or cancer* or tumor* or tumour* or sarcom* or polyp* or adenom*))) AND ((((“occult blood”) AND (faecal or fecal or feces or faeces or stool*)) OR (FOBT* or FIT* or gFOBT*)) AND ((immunohistochem* or immunochem* or immunol* or guaiac*)) OR ((“fecal immunochemical” NEAR/1 test* or “faecal immunochemical” NEAR/1 test* or “fecal immunochemistry” NEAR/1 test* or “faecal immunochemistry” NEAR/1 test* or ColoScreen or Hema-Screen or Hemdetect or Hemoccult or SENSA or Hema-Check or HemaCheck or hemoCARE or Peroheme or ColoCare or Lifeguard or Fecatwin or HemaWipe or Instaccult or Monohaem or Okokit or Seracult or Dencoccult or Early-detector or Earlydetector or Fe-Cult or Fecult or Feca-EIA or FecaEIA or Hemo-FEC or HemoFEC or Hexagon or SureScreen or Hemaprompt or Hemdetect or Camco-PAK or CamcoPAK or Colocheck or Cecogenics or Hematest or Dencocult or Fecatest or Hemofecia or Quick-CULT or QuickCULT))))

Google scholar

“colorectal|colon|colonic|rectal|anal|anus carcinoma|neoplasm|neoplasms|adenocarcinoma|cancer|tumor|tumors” “occult blood” faecal|fecal|feces|faeces|stool|FOBT|FIT|gFOBT

Supplementary Material 2 Variables for which data were extracted

The following data were abstracted when applicable: (i) study characteristics - primary author, journal of publication, year of publication, geographic location of study population, study design (prospective/retrospective), time period of study enrollment, patient selection (inclusion- and exclusion criteria); (ii) FOBT characteristics - type of FOBT used (FIT or gFOBT), brand of FOBT, referral criteria for positive test (i.e. cut-off or number of positive panels), diagnostic test used, diet restrictions; (iii) study cohort characteristics - cohort size, total number of eligible invitees, total number of participants, total tests analyzed, total participants with a positive test, participants demographics (mean age and range, percentage male), reference standard uptake (percentage); (iv) CRC characteristics - total number diagnosed with CRC after negative FOBT, total number diagnosed with CRC after positive FOBT, location of CRC (proximal/distal), CRC stage (I/>I); (v) patient characteristics - gender, age <59 /≥60 years, time of follow-up in years/months (mean, median, min, max), completeness of follow-up (percentage), findings at index diagnostic test.

Supplementary Table 1 Quality assessment of included studies

a) FIT observational studies*

Study	Selection	Comparability	Outcome
Chen[28]	****	*	***
Chiu[29]	****	*	***
Crotta[30]	***	*	**
Digby[33]	****	*	**
Giorgi Rossi[50]	***	*	**
Itoh[35]	***	*	**
Jensen[36]	***	*	***
Launoy[38]	***	*	**
McNamara[41]	***	*	*
Nakama[42]	***	*	***
Parente[44]	***	*	*
Portillo[51]	****	*	***
Shin[46]	***	*	***
van der Vlugt[52]	****	*	***
Zappa[48]	****	*	***

b) gFOBT observational studies*

Study	Selection	Comparability	Outcome
Blom[49]	****	*	***
Bouvier[27]	***	*	***
Cummings[31]	***	*	**
Faivre[34]	****	*	***
Hardcastle[2]	***	*	***
Kewenter[37]	****	*	***
Kronborg[3]	****	*	***
Mandel[4]	****	*	*
Paimela[43]	****	*	**
Rennert[45]	***	*	***
Souques[47]	***	*	*
Steele[6]	****	*	***
Zappa[48]	****	*	***

c) Randomised controlled trials**

Study	Selection bias		Performance bias	Reporting bias	Detection bias	Attrition bias	Other bias
	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Selective reporting	Blinding of outcome assessment	Incomplete outcome data	Anything else, ideally prespecified
Denters[32]	+	+	+	+	+	?	+
Levi[39]	+	+	+	+	+	?	+

* using the Ottawa Newcastle criteria of Wells *et al*.[24]

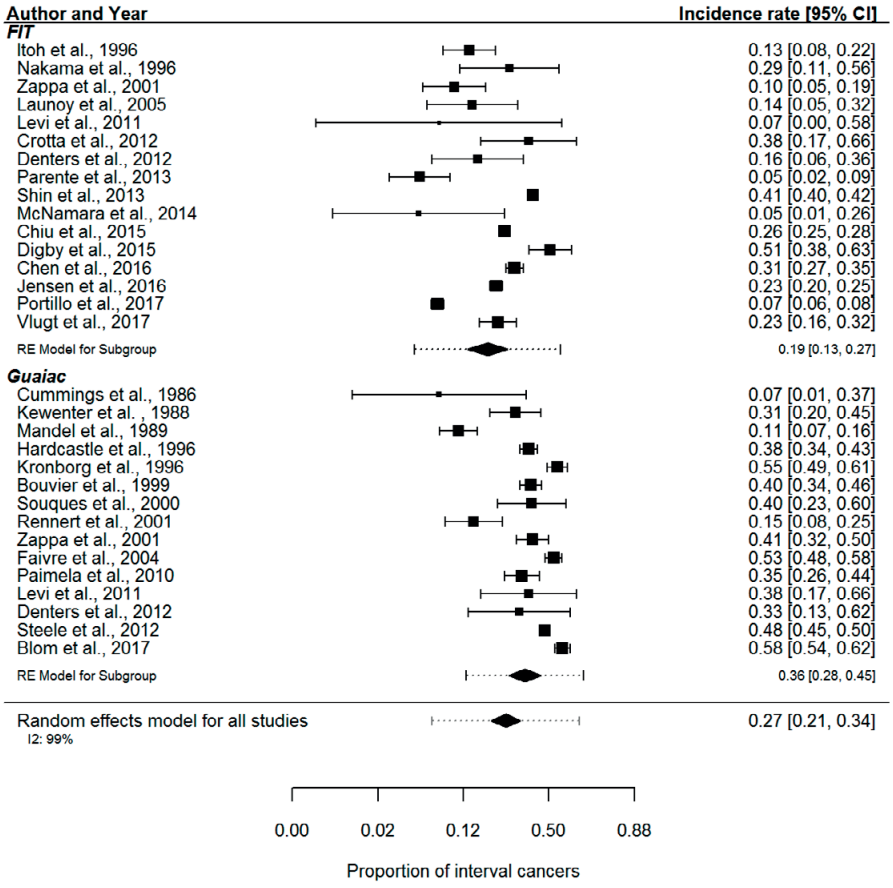
** Both randomised trials were scored as good quality by the Cochrane risk of bias tool.[25] Both studies did not describe handling of incomplete outcome data such as screenees loss to follow-up or missing data when cross-linking the screening pilot database with the cancer registry, therefore, this item was scored as 'unclear risk'. However, this was unlikely to have biased the outcome.

Supplementary Table 2 Quality assessment of the standard and sensitivity analysis of the pooled IRR of FIT iCRC relative to gFOBT iCRC[25]

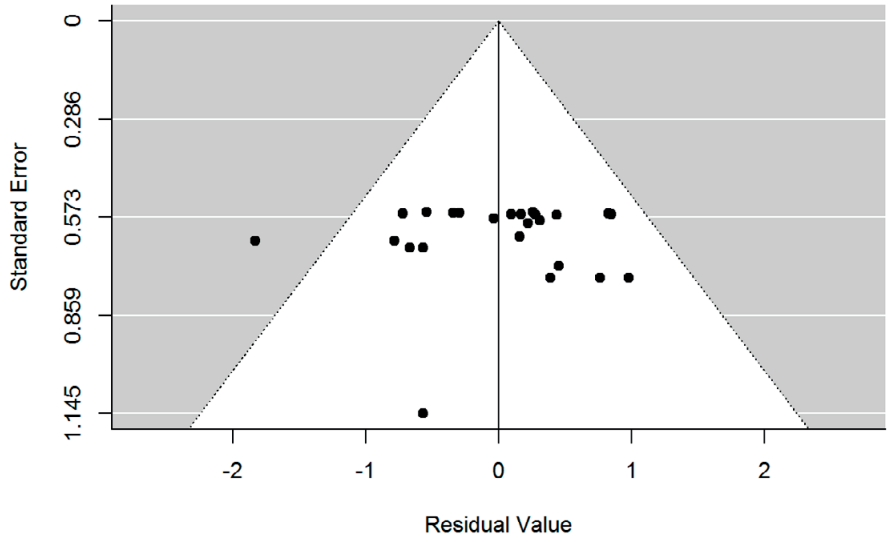
Comparison	**Standard pooled IRR**		**Sensitivity analysis**	
	IRR (95% CI)	Quality of evidence	IRR (95% CI)	Quality of evidence
FIT iCRC versus gFOBT iCRC	0.58 (0.32-1.07)	Very low	0.36 (0·17-0·75)	High

IRR: incidence rate ratio; FIT: faecal immunochemical test; gFOBT guaiac faecal occult blood test; iCRC: interval colorectal cancer

Supplementary Figure 1 Forest plot of FOBT iCRC to screen-detected CRC ratio



Supplementary Figure 2 Funnel plot



Visual inspection of the funnel plot did not show asymmetry and the rank correlation test for asymmetry was not significant (Kendall's tau = -0.0768, p=0.6018).

3

Equivalent accuracy of 2 quantitative fecal immunochemical tests in detecting advanced neoplasia in an organized colorectal cancer screening program

E. Wieten*, C.M. de Klerk*, A. van der Steen, C.R. Ramakers, E.J. Kuipers, B.E. Hansen, I. Lansdorp-Vogelaar, P.M. Bossuyt, E. Dekker, M.C.W. Spaander.

*shared first author

Gastroenterology 2018

ABSTRACT

Background

Although different brands of fecal immunochemical tests (FITs) are used for colorectal cancer (CRC) screening, few studies have compared their accuracy in detecting advanced neoplasia.

Methods

We performed a large prospective cohort study within the Dutch national CRC screening program to evaluate 2 quantitative FITs: FOB-Gold (Sentinel, Italy) and OC-Sensor (Eiken, Japan), from May 2016 through March 2017. We randomly selected 42,179 screening-naïve individuals (55–75 years old), who were asked to perform both FITs themselves using the same bowel movement. Participants with positive results from 1 or both FITs (≥15 μg hemoglobin/gram feces) were invited for colonoscopy examination (reference standard). Equivalence in detection of advanced neoplasia was evaluated with a predefined margin of 0.15%.

Results

Of 42,179 invitees, 22,064 (52%) participated and FITs were completed for 21,078 participants. Of 2112 participants (9.6%) with 1 or 2 positive results from FITs, 1778 (84%) underwent a colonoscopy. Of all invitees, the FOB-Gold test detected advanced neoplasia (confirmed by colonoscopy) in 610 participants (1.45%) and the OC-Sensor detected advanced neoplasia (confirmed by colonoscopy) in 606 participants (1.44%)-an absolute difference of 0.01% (95% confidence interval [CI], –0.06% to 0.08%). Of the 21,078 participants who completed both FITs, 1582 (7.5%) had a positive result from the FOB-Gold test and 1627 (7.7%) a positive result from the OC-Sensor test (*P*=.140). The relative true-positive rate of FOB-Gold vs OC-Sensor in detecting advanced neoplasia was 0.97 (95% CI, 0.92–1.01) and 0.95 (95% CI, 0.87–1.03) for CRC. The relative false-positive rate of the FOB-Gold test vs the OC-Sensor test in detecting advanced neoplasia was 0.99 (95% CI, 0.93–1.05).

Conclusion

In a large prospective study of individuals invited for CRC screening in The Netherlands, we found equivalent accuracy of the FOB-Gold FIT vs the OC-Sensor FIT in detecting advanced neoplasia. These results are relevant for selecting FITs for CRC screening programs worldwide. Dutch National Trial Registry: NTR5874.

INTRODUCTION

Colorectal cancer (CRC) is a significant public health problem due to its high incidence, morbidity and mortality.[1] It is the third most commonly diagnosed cancer in men and second in women.[1] CRC screening by means of guaiac fecal occult blood testing (gFOBT) and sigmoidoscopy has been shown to reduce CRC-related mortality through the detection of advanced adenomas and early-stage CRC.[2-6] Accordingly, CRC screening programs are being implemented worldwide.[7]

Fecal immunochemical tests (FITs) are rapidly replacing gFOBTs for CRC screening, because of their easier handling, higher uptake, automated assessment, and higher sensitivity in detecting advanced neoplasia (AN).[8] In addition, FITs can provide a quantitative test result, allowing cutoff adjustment to match available colonoscopy resources. FITs also allow for single stool testing, and do not require dietary restrictions, and, consequently, lead to higher participation rates than gFOBTs.[8-11] As a consequence, many countries are either in the transition from gFOBT to FIT screening, or implement FIT screening from the start.[7]

In the European guidelines, the quantitative immunochemical tests are recommended as test of choice for population screening.[9] These guidelines also indicate that a screening program should assess individual device characteristics, including accuracy, ease of use by participant and laboratory, suitability for transport, sampling reproducibility and sample stability. However, comparative data with enough power from head-to-head comparisons to decide on equivalence of FIT assays in detecting AN in population screening are not yet available, hindering informed decision making.

At present, multiple FITs are available, of which FOB-Gold (Sentinel, Milan, Italy) and OC-Sensor (Eiken Chemical, Tokyo, Japan) are currently widely used.[12] FIT assays vary in analytical performance due to a range of factors, including anti-heme antibody characteristics and assay optimization, buffer composition and volume and sample tube design. These differences influence the measured fecal hemoglobin (Hb) concentration, FIT positivity rate, error rates, and capacity to detect AN.[13, 14] One randomized controlled trial reported higher true-positivity and false-positivity rates for FOB-Gold compared to OC-Sensor at equal positivity cutoffs (in ng Hb/mL buffer).[15] Three other randomized trials reported equal AN detection rates of both tests, but higher positivity rates with FOB-Gold.[13, 16, 17] However, all these studies were small and randomized invitees to one of both tests.[13, 15, 16] A recent comparative evaluation of 9 FITs, including OC-Sensor and FOB-Gold, showed that the sensitivity to detect AN varied widely between FITs, ranging from 16% to 34%.[18] Yet, this retrospective study was performed in a laboratory setting using stored samples and limited inclusions per test. A large population-based study with a paired design, in which both tests are compared within the same individual and sampled

from the same stool would minimize the risk of confounding factors and increase the applicability of the study results to CRC screening programs.

We therefore conducted a large prospective population-based study within the Dutch nationwide CRC screening program to compare the accuracy in detecting AN for 2 widely used FITs.

METHODS

Study Population

This study was conducted within the Dutch CRC screening program between May 2016 and March 2017. The structure of the Dutch CRC screening program has been described elsewhere.[19] In short, demographic data of all individuals between the ages of 55 and 75 years, living in the southwest region of the Netherlands, were obtained from municipal registers. The target population for our study consisted of first-time invitees in 2016, that is those aged 59, 61, 63, 71, or 75 years.[19] This population encompassed more than 250,000 eligible screening-naïve individuals. For the purpose of this study, a random sample was taken from this target population with a computer run algorithm (SPSS, version 23, IBM Corp, Armonk, NY). Selection of study invitees preceded invitation. This study was ethically approved by the Minister of Health (Population Screening Act; no. 769500-1357 16-PG).

Study design and intervention

Two FITs (one FOB-Gold and one OC-Sensor) were sent by postal mail. Individuals were invited to collect a single feces sample of the same bowel movement with each test. Detailed sampling instructions as recommended by the manufacturer were provided for each test. Study invitees were asked to fill out an informed consent form, including the sampling date. After 42 days, a reminder was sent automatically to nonresponders. Consenting invitees were asked to return both tests and consent form in a prepaid and sealed envelope to one accredited, centralized laboratory. Persons who had actively deregistered from the screening program, who had moved, were deceased, did not consent or did not respond to the study invitation were labelled as nonparticipants.[19]

Supplementary Table 1a and 1b describe the pre-analytical aspects and analytical performance of the FIT analyses. The OC-Sensor tests were analyzed by using the OC-Sensor Diana analyzer. The FOB-Gold tests were analyzed by using the Bio Majesty JCA-BM6010/C analyzer. Quantitative results for both tests were provided in ng Hb/mL. For the purpose of the study, FOB-Gold test were considered positive at ≥88 ng Hb/mL and OC-Sensor tests at ≥75 ng Hb/mL. At the time that the study was designed, a positivity cut-off of ≥88 ng Hb/mL was used in the Dutch nationwide CRC screening program for FOB-Gold tests. To be able to compare FOB-Gold to OC-Sensor tests within the nationwide

CRC screening program, OC-Sensor test positivity cut-off (in µg Hb/g feces) was set to be the same as that for FOB-Gold. Converted into micrograms (µg) of Hb per gram of stool, the threshold for a positive test result was ≥15 µg Hb/g feces for both tests.[20]

Participants were referred for colonoscopy in case of 1 or 2 positive FIT results. Participants with 2 negative test results were referred back to the Dutch CRC screening program and will be re-invited after 2 years. Participants who returned 2 non-analyzable tests, who had 2 unreliable test results, or for whom both tests were missing, were referred back to the Dutch CRC screening program and were not included for the primary outcome of this study.

Participants were informed about their FIT results by postal mail within 5 days after the FIT was analyzed. If 1 or both FIT results were positive, the family physician was informed and the participant was invited for a precolonoscopy interview in an accredited colonoscopy center nearby the participant's home address. At this precolonoscopy interview, participants' eligibility for colonoscopy was assessed. Colonoscopy exclusion criteria were similar to those of the Dutch CRC screening program: a life expectancy of 5 years or less, a proctocolectomy in the past, under current treatment for CRC, history of inflammatory bowel disease, and a complete colonoscopy in the past 5 years.[21] Colonoscopies were performed within 10 days after the interview at one of the certified colonoscopy centers by accredited endoscopists who performed at least 200 colonoscopies a year with an adenoma detection rate of ≥ 30%.

Location, size and morphology of all identified colorectal lesions were reported using an automated structured colonoscopy reporting system. Polyps were removed and sent for pathology review in separate jars.[22] Advanced adenoma was defined as an adenoma ≥ 10 mm, with ≥ 25% villous component and/or high-grade dysplasia. AN included advanced adenoma or CRC. If multiple lesions were present, the participant was classified according to the most advanced lesion. Logistics were executed conform the Dutch CRC screening quality guidelines.[19 23] Socioeconomic status was assessed using the Dutch area social status score, grouped into quintiles. These scores are developed by the Netherlands Institute of Social Research, and are a composite measure of education, income and employment.[24]

Outcome measures and statistical analysis

Our primary outcome measure was the difference in diagnostic yield of AN between OC-Sensor and FOB-Gold, defined as the number of participants with AN detected relative to the number of invitees.

We hypothesized that the 2 FITs would generate an equivalent diagnostic yield, defined as an absolute symmetric difference of at most 0.15%, relative to the number of invitees. Differences in paired proportions of diagnostic yield were evaluated using the method described by Liu et al.[25] Confidence intervals (CI) were calculated using a Wald interval with Bonett-Price adjustment.[26] Comparisons in diagnostic yield for CRC between both tests were evaluated similarly.

In addition, we compared the accuracy of both FITs in participants with paired test results, that is, in participants with 2 complete and reliable FIT results. Excluded from paired analysis were participants with 1 or 2 non-analyzable tests (due to fecal overload, loss of two-thirds or more of the total buffer volume both from visual assess by laboratory staff and by automatic system, missing barcode or another technical problem) and participants with an unreliable test result (in case the return date was more than 6 days after sampling or if the sampling date was missing). Paired accuracy was assessed by calculating the relative true-positive rate and the relative false-positive rate. The relative true-positive rate is defined as the number of true-positive results (in whom AN was detected at colonoscopy) for FOB-Gold relative to the number of true-positive results for OC-Sensor. The relative true-positive rate is equal to the relative sensitivity rate in screening participants. A relative sensitivity of 1.00 implies that both screening tests result in the same sensitivity, a relative sensitivity of 1.10 would imply that using FOB-Gold results in 10% more true-positives. The relative true-positive rate was calculated with 95% CIs using the methods proposed by Alonzo et al.[27] Similarly, the relative false-positive rate was estimated, defined as the number of false-positive test results (in whom no AN was detected at colonoscopy) for FOB-Gold relative to the number of false-positive test results for OC-Sensor.[27]

Our secondary outcomes included the number of analyzable tests, the positivity rate (defined as the proportion of participants with a positive FIT result) and the positive predictive values (defined as the number of participants in whom AN was detected relative to those undergoing colonoscopy after a positive FIT result). The participation rate was calculated as the number of participants relative to the number of invitees.

Paired proportions were compared using the McNemar test. Other proportions were compared using chi-square statistics. All p values were 2-sided; differences were considered significant if $p<0.05$.

Sample size calculations

In this equivalence trial, we assumed a 1.5% diagnostic yield for AN with both tests, anticipated no difference, and wanted to exclude an absolute difference in diagnostic yield of 0.15% or more between the 2 tests. We expected that 5 per 1000 invitees would

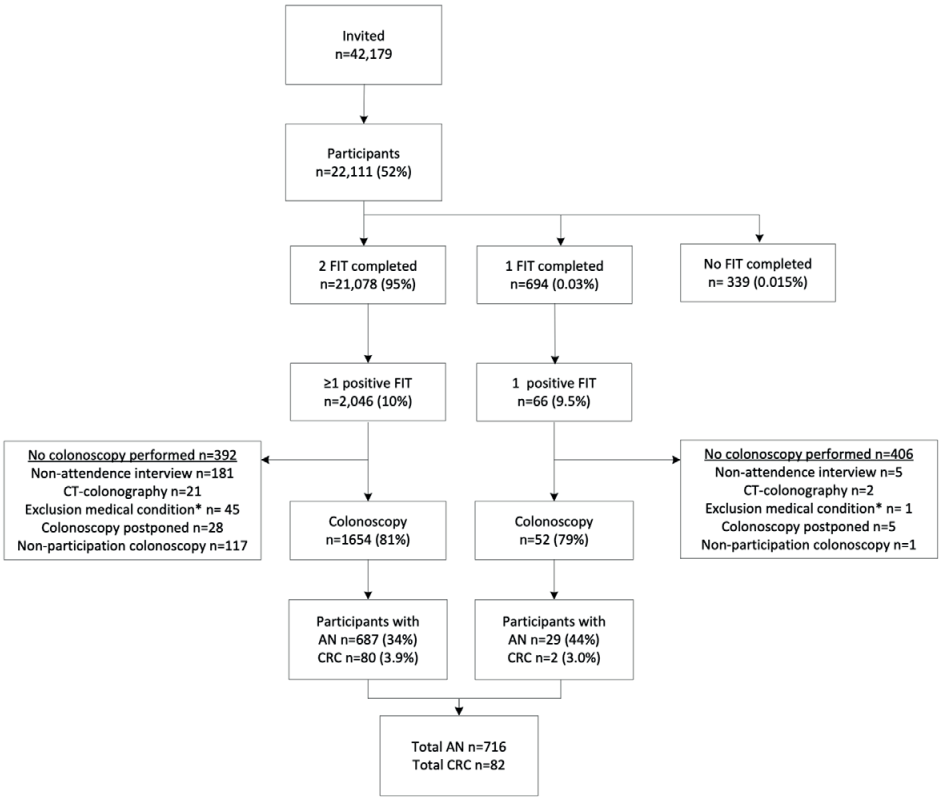
have discordant true-positive FIT results: invitees with AN detected after 1 (not 2) positive FIT results. In that case, inviting 40,000 screenees, with an expected participation rate of 50% (20,000 screenees with 2 complete test returns), would give us more than 90% power to demonstrate equivalence, using an alpha of 0.05.[25] All authors had access to the study data and reviewed and approved the final manuscript.

RESULTS

Cohort characteristics

Of 42,179 individuals invited for screening, 22,064 (52%) participated in the study (Figure 1). Baseline characteristics of study participants are shown in Table 1. Of the participants, 21,078 (96%) completed both FITs in a single bowel movement, 694 (3.2%) completed only one FIT, and 292 returned two incomplete tests (non-analyzable tests or tests with unreliable results).

Figure 1 Study flow



FIT: fecal immunochemical test; CRC: colorectal cancer; AN: advanced neoplasia; CT computer-tomography

Table 1 Characteristics of study participants at baseline, with ≥1 positive fecal immunochemical test results, and at colonoscopy

Total	Participants	Participants with ≥1 positive FIT result	Participants with colonoscopy
	n=22,057	n=2112	n=1778
Male sex, n (%)	10,589 (50)	1256 (60)	1066 (60)
Age*, median (IQR)	60 (58-62)	60 (59-63)	60 (59-63)
Socioeconomic status, n (%)			
- Very high	4407 (20)	397 (19)	341 (19)
- High	4843 (22)	387 (18)	330 (19)
- Average	4052 (18)	372 (18)	316 (18)
- Low	5136 (23)	536 (25)	451(25)
- Very low	3577 (16)	416 (20)	336 (19)
- Missing	49 (<1)	4 (<1)	4 (<1)

*Age at time of FIT invitation
IQR=interquartile range
Socioeconomic status scores are a composite area-based measure of education, income and employment.

All 2112 (9.6%) with 1 or 2 positive FIT results were invited for a precolonoscopy interview; 1778 (84%) underwent a colonoscopy. Of those who attended colonoscopy 1066 were male (60%) and median age was 60 years (interquartile range 59-63 years; Table 1). In 716 (4.0%) participants with a positive FIT result who underwent a colonoscopy AN was detected and in 82 (0.5%) CRC was detected.

Diagnostic yield for advanced neoplasia and colorectal cancer

AN was detected after a positive FOB-Gold in 610 of total invitees (1.45%) and in 606 (1.44%) after a positive OC-Sensor, resulting in an absolute difference of 0.01% (95%CI -0.06% to 0.08%). The existence of a difference of 0.15% or more between FOB-Gold and OC-Sensor in detection of AN was rejected ($p < 0.001$). CRC was detected after a positive FOB-Gold result in 74 invitees (0.18%) and after a positive OC-Sensor result in 78 (0.18%), resulting in an absolute difference of -0.009% (95%CI -0.027% to 0.008%).

Incomplete test returns and non-analyzable tests

Differences between FOB-Gold and OC-Sensor tubes in the proportion of non-analyzable results are shown in Table 2. Forty-nine of the 22,057 (0.22%) returned FOB-Gold tests were non-analyzable vs 14 of 21,369 returned OC-Sensor tests (0.07%; $p<0.001$). Buffer loss was more frequently a problem for analysis of FOB-Gold tests than for OC-Sensor ($p=0.005$).

Paired accuracy: relative true-positive rate and relative false-positive rate

The results of 986 (4.5%) participants were not included in the paired analysis due to not returning 1 of the tests (n=702), ≥ 1 non-analyzable test(s) (n=54) or ≥ 1 unreliable test

Table 2 Reasons for non-analyzable fecal immunochemical test tubes

Total returned tests	FOB-Gold n=22,057	OC-Sensor n=21,369	P-value
Non-analysable tests, n (%)	**49 (0.22)**	**14 (0.07)**	**<0.001**
Technically impossible	10 (20)	4 (29)	0.52
Barcode unreadable	5 (10)	4 (29)	0.09
Broken tube	2 (4.1)	0	0.45
No buffer	23 (47)	1 (7..1)	0.005
Too large sample	6 (12)	1 (7.1)	0.60
Too small sample	3 (6.1)	3 (21)	0.09
No sample taken	0	1 (7.1)	0.06

result(s) (n=260), or a combination of these (n=30). 695 participants did not provide a OC-Sensor test result and 7 an FOB-Gold test result. Of the 21,078 participants who completed both FITs, the positivity rate was 7.5% for FOB-Gold compared to 7.7% for OC-sensor (p=0.140) (Table 3). A total of 1163 (57%) had 2 positive FITs, 419 (20%) participants had a positive FOB-Gold and negative OC-Sensor result, and 464 (23%) participants a negative FOB-Gold and positive OC-Sensor result.

The relative true-positive rate (relative sensitivity) for the detection of AN for FOB-Gold relative to OC-Sensor was 0.97 (95%CI: 0.92 to 1.01; Figure 2), and the false-positive rate was 0.99 (95%CI: 0.93 to 1.05). Positive predictive value of AN was 43.8% for FOB-Gold and 44.3% for OC-sensor (absolute difference 0.5%; 95%CI: -3.3% to 4.2%, p <0.05). The

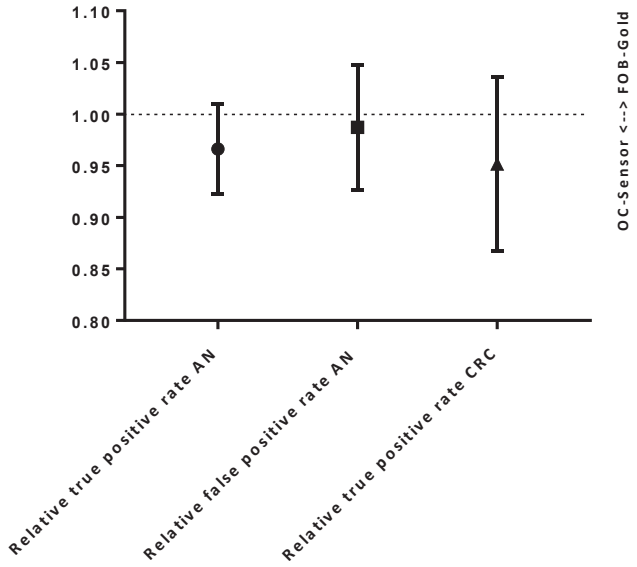
Table 3 Detection of advanced neoplasia and colorectal cancer with FOB-Gold and OC-Sensor in paired design

		FOB-Gold		
		Positive*	**Negative**	**Total**
OC-Sensor	**Positive***	1163	464	1627 (7.7%)
	AN	500	104	604 (2.9%)
	CRC	70	7	77 (0.4%)
	Negative	419	19,032	19,451 (92.3%)
	AN	83	-	-
	CRC	3	-	-
	Total	1582 (7.5%)	19,496 (92.5%)	21,078
	AN	583 (2.8%)	-	-
	CRC	73 (0.4%)	-	-

Results for participants with two completed FITs only
AN: advanced neoplasia (CRC and/or advanced adenoma, defined as an adenoma ≥ 10 mm, with ≥ 25% villous component and/or high-grade dysplasia).
CRC: colorectal cancer
*Overall colonoscopy attendance rate of participants with ≥positive FIT(s) was 84% at fecal haemoglobin (Hb) positivity cut-off of ≥ 15 µg/Hb feces.

relative true-positive rate (relative sensitivity) for FOB-Gold versus OC-Sensor in detecting CRC was 0.95 (95%CI: 0.87 to 1.03; Figure 2). Positive predictive value of CRC was 5.5% for FOB-Gold and 5.7% for OC-sensor (absolute difference 0.2%; 95%CI: -1.5% to 2.0%).

Figure 2 Relative true-positive and false-positive rates FOB-Gold vs OC-Sensor with 95% confidence intervals



CRC: colorectal cancer; AN: advanced neoplasia

DISCUSSION

This large prospective trial within the Dutch national CRC screening program shows that 2 widely used FITs, FOB-Gold and OC-Sensor, have similar accuracy in detecting AN and CRC at a positivity cut-off level of ≥15 µg Hb/g feces.

Factors adding to the validity of our results include the study's paired design, its implementation within the logistics of the nationwide screening program, and its large and representative cohort of participants. Moreover, both FITs were analyzed in the same laboratory at the day of arrival and followed identical logistic routes. Consequently, factors that are known to influence test results were identical, including temperature changes, time differences from sampling to analyzing, and laboratorial logistic differences.[28] Screening logistics, from invitation to pathology for those who underwent colonoscopy, were identical to those of the Dutch nationwide screening program.

Some potential limitations also have to be acknowledged. Participation rate in this study was lower (52%) than in the current Dutch nationwide screening program (72%) and previously performed pilot studies in our country (60-63%).[13, 29,] We had anticipated this difference since an alternative existed for our study invitees: those who opted out of the study could still participate in the regular one-FIT population CRC screening program. For this reason, we invited over 40,000 persons for the study, to have enough power and precision for a comparison of the 2 FITs. As no colonoscopies were performed in participants with 2 negative test results, we are unable to provide estimates of sensitivity, yet we calculated estimates of relative sensitivity. We also observed that more OC-Sensor tubes were missing than FOB-Gold tubes. This was most likely due to the way the study was embedded within the nationwide screening program. Study invitees were asked to perform both tests if they wished to participate in the study. Of those invitees that did not want to participate in the study but had a preference to participate in the regular nationwide screening program with 1 test, only the FOB-Gold was analyzed in the laboratory. The form was checked twice, first manually by laboratory assistants and subsequently by an automated system. In case no study approval was given by the screenee, only the FOB-Gold was analyzed by the laboratory assistant as this is the test that is currently used in the nationwide CRC screening program. However, this means some OC-Sensor tests may have been falsely qualified as being not returned by the screenee, while in reality the test was sent back with written study consent. This could have occurred when the manual check for informed consent by the laboratory staff differed from the scanned consent by the computer.

Previous studies comparing FIT brands in CRC screening studied either randomized subjects to perform a single FIT [13, 15-17], or had FITs performed in different bowel movements.[30] Differences in true-positive rate in favor of OC-Sensor to FOB-Gold were found in a Spanish screening cohort, in which participants were randomly invited to perform one of both tests.[15] The cut-off was equalized in ng Hb/mL buffer, instead of µg Hb/g feces, therefore subject to known differences in buffer volumes between both FIT brands (1.7 mL FOB-Gold and 2.0 mL OC-Sensor).[20] Another randomized trial in France showed equal detection rates for both FITs but found a lower positive predictive value for FOB-Gold than OC-Sensor.[17] In a Latvian screening cohort, no difference between FOB-Gold and OC-Sensor was found in number needed to scope to detect 1 participant with AN at a positivity cut-off of 15 µg Hb/g faeces.[16] In a Dutch screening pilot, diagnostic yield and positive predictive value were equal for both FITs at equal cutoffs, and equal positivity rates.[13] The authors recommended to compare FIT brands at equal positivity rates, rather than equal cutoffs. In our study, the same cutoff yielded comparable positivity rates. In contrast to our study, these previous studies were not powered to determine small differences in the detection of AN.

Almost half of the participants with a positive test result had only 1 positive FIT result. One could wonder whether such discordant results will also be detected when 2 FITs of the same brand are compared. Discordant detection and positivity rates between 2 identical FITs might result from an uneven distribution of hemoglobin through the feces. It should also be noted that the sampling instructions by the manufacturers of FOB-Gold and OC-Sensor differ, which may influence test handling and, possibly, its results. FOB-Gold instructions prescribe inserting the sampling stick in 4 different parts of the stool. The OC-Sensor on the other hand, should be scraped through the stool in 4 different parts. Both sets of instructions were provided as such to our study invitees. We found, however, no evidence that these sample methods resulted in different detection rates.

In line with existing evidence, we found more sampling errors with the FOB-Gold than OC-sensor, although proportions in our study were smaller (0.22% vs 0.07%) than in the Dutch pilot screening program (2.0% vs 0.7%, $p<0.001$)[13], the Spanish study (2.3% vs 0.2%)[15], and within a Latvian screening setting (4.4% vs 0.2%, $p<0.001$)[31]. This might be explained by the fact that FOB-Gold adjusted its cap to prevent buffer loss by opening the wrong side of the tube. Furthermore, in the Dutch pilot screening program, screenees were probably more used to the OC-Sensor test because this test was used in the first 3 rounds and FOB-Gold was only introduced in the fourth. We found, however, no evidence that these sample methods resulted in different detection rates.

CRC is the second most common cancer and cause of cancer-related death worldwide.[1] To reduce CRC incidence and mortality in Europe, the European Committee have made recommendations on CRC screening for all European Union members.[32] They stated that the desired level of screening coverage of the target population is 95% with a desirable level of uptake of >65%.To fulfil these target levels, one needs to have an organized population-based screening program with a high participation rate for the selected screening tool. Participation rates of FIT are generally high, therefore FIT has been selected in most countries as the screening test of choice.[7]

As we found no evidence that FOB-Gold and OC-Sensor differ in detecting AN and CRC, other features can now guide informed decision making when selecting 1 of these 2 brands for FIT-based CRC screening. Probably the next most important consideration is ease of use of the test and its effect on participation rate. As stated by Winawer and Allison: 'the best test is the one that gets done well'.[33] Many other factors should be considered including costs, ease of use for laboratory staff or other stakeholders involved in FIT analysis, suitability for transport, the keeping quality of the tubes, analyzer features, capacity, speed, analytical performance, sample stability, easy of handling, safety during

postage and labeling. Depending on context and setting, more studies are warranted to evaluate these other aspects of FIT.

When this trial started, the Dutch nationwide screening program was already launched. Yet evaluations of screening strategies should not be limited to the implementation phase. As others have argued, evaluation should continuously be explored within and while running a screening program, with formal study designs such as paired comparisons or randomized trials.[34] Apart from its results, our study may therefore serve as an example how to assess and possibly improve screening effectiveness within an ongoing program and may, hopefully, inspire future initiatives.

Acknowledgements
The authors would like to thank the Netherlands Organization for Health Research and Development of the Dutch Ministry of Health (ZonMw) for funding, and all participants for their efforts to perform two tests and a questionnaire. The authors thank all involved co-workers of the Foundation of Population Screening South-West (Bevolkingsonderzoek Zuid-West), and all involved co-workers at Daklapack packaging for their contributions to the study. Finally, the authors thank Star lab (Rotterdam, The Netherlands), for their efforts in analyzing twice as many FITs compared to the nationwide Dutch screening program during the study period.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359-86.
2. Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med. 2012;366(8):687-96.
3. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-77.
4. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev. 2007(1):CD001216.
5. 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.
6. Holme O, Loberg M, Kalager M, Bretthauer M, Hernan MA, Aas E, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. JAMA. 2014;312(6):606-15.
7. 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.
8. Kuipers EJ, Spaander MC. Colorectal Cancer Screening by Colonoscopy, CT-Colonography, or Fecal Immunochemical Test. J Natl Cancer Inst. 2016;108(2):djv383.
9. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
10. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol. 2013;10(3):130-42.
11. Digby J, McDonald PJ, Strachan JA, Libby G, Steele RJ, Fraser CG. Use of a faecal immunochemical test narrows current gaps in uptake for sex, age and deprivation in a bowel cancer screening programme. J Med Screen. 2013;20(2):80-5.
12. Schreuders EH, Grobbee EJ, Spaander MCW, Kuipers EJ. Advances in Fecal Tests for Colorectal Cancer Screening. Current Treatment Options in Gastroenterology. 2016;14(1):152-62.
13. 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. 2016;66(11):1975-82.
14. Chiang TH, Chuang SL, Chen SL, Chiu HM, Yen AM, Chiu SY, et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cutoff concentration in a nationwide colorectal cancer screening program. Gastroenterology. 2014;147(6):1317-26.
15. Zubero MB, Arana-Arri E, Pijoan JI, Portillo I, Idigoras I, Lopez-Urrutia A, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test. Front Pharmacol. 2014;4:175.
16. Santare D, Kojalo I, Liepniece-Karele I, Kikuste I, Tolmanis I, Polaka I, et al. Comparison of the yield from two faecal immunochemical tests at identical cutoff concentrations - a randomized trial in Latvia. Eur J Gastroenterol Hepatol. 2016;28(8):904-10.

17. Faivre J, Dancourt V, Denis B, Dorval E, Piette C, Perrin P, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer. 2012;48(16):2969-76.
18. Gies A, Cuk K, Schrotz-King P, Brenner H. Direct Comparison of Diagnostic Performance of 9 Quantitative Fecal Immunochemical Tests for Colorectal Cancer Screening. Gastroenterology. 2018;154(1):93-104.
19. 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. 2017;152(4):767-75.e2.
20. Fraser CG, Allison JE, Halloran SP, Young GP, Org WE. A Proposal to Standardize Reporting Units for Fecal Immunochemical Tests for Hemoglobin. Jnci-J Natl Cancer I. 2012;104(11):810-14.
21. National Institute for Public Health and the Environment (RIVM). Exclusion criteria colonoscopy (Dutch) http://www.rivm.nl/Documenten_en_publicaties/Professioneel_Praktisch/Richtlijnen/Preventie_Ziekte_Zorg/Darmkanker/Exclusiecriteria_coloscopie (date accessed 20th June 2017); 2013.
22. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol. 2007;29(1):19-24.
23. National Institute for Public Health and the Environment (RIVM). Uitvoeringskader Bevolkingsonderzoek Darmkanker (Dutch). www.bevolkingsonderzoekdarmkanker.nl/downloads (date accessed 20th June 2017); 2017.
24. 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.
25. Liu JP, Hsueh HM, Hsieh E, Chen JJ. Tests for equivalence or non-inferiority for paired binary data. Stat Med. 2002;21(2):231-45.
26. Fagerland MW, Lydersen S, Laake P. Recommended tests and confidence intervals for paired binomial proportions. Stat Med. 2014;33(16):2850-75.
27. Alonzo TA, Braun TM, Moskowitz CS. Small sample estimation of relative accuracy for binary screening tests. Stat Med. 2004;23(1):21-34.
28. De Girolamo G, Goldoni CA, Corradini R, Giuliani O, Falcini F, Sassoli De'Bianchi P, et al. Ambient temperature and FIT performance in the Emilia-Romagna colorectal cancer screening programme. J Med Screen. 2016;23(4):186-91.
29. Bevolkingsonderzoek darmkanker: Monitor 2016 (Dutch). http://www.rivm.nl/Documenten_en_publicaties/Algemeen_Actueel/Uitgaven/Preventie_Ziekte_Zorg/Darmkankerscreening/Landelijke_Monitoring_en_Evaluatie_Bevolkingsonderzoek_Darmkanker.
30. Kapidzic A, van Roon AH, van Leerdam ME, van Vuuren AJ, van Ballegooijen M, Lansdorp-Vogelaar I, et al. Attendance and diagnostic yield of repeated two-sample faecal immunochemical test screening for colorectal cancer. Gut. 2017;66(1):118-23.
31. Santare D, Kojalo I, Huttunen T, Rikacovs S, Rucevskis P, Boka V, et al. Improving uptake of screening for colorectal cancer: a study on invitation strategies and different test kit use. Eur J Gastroenterol Hepatol. 2015;27(5):536-43.
32. Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). Off J Eur Union. 2003:34-8.
33. Allison JE. The best screening test for colorectal cancer is the one that gets done well. Gastrointest Endosc. 2010;71(2):342-5.

34. Bell KJ, Bossuyt P, Glasziou P, Irwig L. Assessment of changes to screening programmes: why randomisation is important. BMJ. 2015;350:h1566.
35. Fraser CG, Allison JE, Young GP, Halloran SP, Seaman HE. Improving the reporting of evaluations of faecal immunochemical tests for haemoglobin: the FITTER standard and checklist. Eur J Cancer Prev. 2015;24(1):24-6.

Supplementeray Table 1a FOB-Gold Standard for Fecal Immunochemical Tests for hemoglobin (FITTER) checklist[35]

Topic	Item	Documentation
Specimen collection and handling		
	Name of specimen collection device and supplier (address).	Name collection device: FOB-Gold Supplier: FOB-Gold, Sentinel Diagnostics, Via Robert Koch, 2 - 20152 Milan – Italy.
	Description of specimen collection device (vial with probe/stick, card, other).	Round tube with collection stick immerged in a preservative solution
	Description of specimens used if an *in vivo* study (single or pooled faeces, artificial matrix with added blood, etc).	Single human faeces sample.
	Details of fecal collection method (sampling technique and number of samples).	Ribbed section of the sampling stick is dipped in four different parts of the stool.
	Who collected the specimens from the samples (patient, technician, etc).	Participant
	Number of fecal specimens used in the study (single, pooled, individual patient faeces).	Single sample of individual patient faeces
	Mean mass of faeces collected.*	10 mg
	Volume of buffer into which specimen is taken by probe, applicator stick or card.*	1.7 mL
	Time and storage conditions of fecal specimen from "passing" to sampling, including time and temperature (median and range).	Analysis took place at same day of arrival (<24 hours) of the FIT in the lab and the FIT was kept by ambient air temperature.
	Time and storage of collection devices from specimen collection to analysis, including time and temperature (median and range). A concise description of process from collection to analysis is recommended.	Participants were asked to post the faeces samples within 24 hours after collection and keep the sample in the refrigerator. The date of sample collection is noted. FIT was transported and analyzed by ambient air temperature. All samples were analyzed within 5 days after collection.
Analysis		
	Name of analyzer, model, supplier (address), number of systems if more than one used.	Bio Majesty JCA-BM6010/C, serial number CA 1401000690069. Supplier: Sysmex Nederland B.V. Ecustraat 11 4879NP Etten Leur
	Number of times each sample was analyzed.	Single or twice. If first analysis resulted in 'no results' analysis was repeated.

Analytical working range* and whether samples outside this range were diluted (factor) and re-assayed.	0.4-797.2 ng/mL. Client samples outside this linearity range were not diluted.
Source of calibrator(s) (supplier with address), number of calibrator(s), how concentrations were assigned* and details of calibration process including frequency.	Calibrator supplier: Sentinel Diagnostics, Via Robert Koch, 2 - 20152 Milan - Italy. Calibrator levels: 6 Standard calibration is performed with every reagent and calibrator lot number change.
Analytical imprecision*, ideally with number of samples analyzed, concentrations, and mean, SD and CV.	Prior to the go-live a CLSI EP5A2 protocol was performed on all three Sentinel controls (low-mid-high) to verify the imprecision specifications as stated in the tender requirements of the colon cancer screening program. CLSI EP5A2 results: Sentinel LOW 50 ng/mL Lot number control 30004/A0546 SD with-in (calculated) = 4.04 SD with-in (claim) = 5.00 (10% of 50 ng/mL) (User variance/claim variance)*freedom degrees = 26.13 Critical Chi-square value = 55.76 **Claim accepted? YES** SD total (calculated) = 5.12 SD total (claim) = 7.50 (15% of 50 ng/mL) (User variance/claim variance)*freedom degrees = 25.63 Critical Chi-square value = 73.03 **Claim accepted? YES** Sentinel MID 71 ng/mL Lot number control 30004/A0551 SD with-in (calculated) = 3.69 SD with-in (claim) = 7.10 (10% of 71 ng/mL) (User variance/claim variance)*freedom degrees = 10.82 Critical Chi-square value = 55.76 **Claim accepted? YES** SD total (calculated) = 5.49 SD total (claim) = 10.60 (15% of 71 ng/mL) (User variance/claim variance)*freedom degrees = 11.98 Critical Chi-square value = 61.66 **Claim accepted? YES** Sentinel HIGH 312 ng/mL Lot number control 30004/A0552 SD with-in (calculated) = 4.81 SD with-in (claim) = 31.20 (10% of 312 ng/mL) (User variance/claim variance)*freedom degrees = 0.95 Critical Chi-square value = 55.76 **Claim accepted? YES** SD total (calculated) = 7.41 SD total (claim) = 46.80 (15% of 312 ng/mL) (User variance/claim variance)*freedom degrees = 1.08 Critical Chi-square value = 59.30 **Claim accepted? YES**

Quality management		
	Source (address) or description of internal quality control materials, number of controls, assigned target concentrations and ranges, how target concentrations were assigned, rules used for acceptance and rejection of analytical runs.	3 rounds of control before running daily analyses were done and 3 rounds after, conform Sentinel's quality rules. If 2 out of 3 controls are within the range, analytical runs are accepted. Apart from the Sentinel controls, a mid-daily run of control conform SKML (SKML CFB, Mercator 1, Toernooiveld 214, NL-6525 EC, Nijmegen, The Netherlands) is performed, every other day a high run or low run: SKML low: 212 ng/mL ->5.05% SKML high: 510 ng/mL ->3.10% If the control is not right, controls are being repeated, if not right after multiple control rounds, de clinical chemist is consulted.
	Participation in external quality assessment schemes: (name and address of scheme), frequency of challenges, performance attained.	Participation in external quality assessments of SKML (foundation of quality control of medical laboratory diagnostics) following a fixed schedule. Assessment results are monitored by the national functionary iFOBT.
	Accreditation held by the analytical facility (address).	Accreditation by CCKL, Mariaplaats 21-D, 3511 LK UTRECHT
	The number, training and expertise of the persons performing the analyses and recording the results	7 trained technician's
Result handling		
	Mode of collection of data-manual recording or via automatic download to IT system, single or double reading	Results are automatically uploaded to the Colonissystem, after authorization by the laboratory analyst the results are uploaded to the screening IT system ScreenIT.
	Units used, with conversions to µg Hb/g faeces if ng Hb/mL used.	In analyzing and reporting results ng Hb/mL was used. For reporting in publications this is converted to µg Hb/g faeces.[200]
	Cut-off concentration(s) if used and explanation of how assigned locally or by manufacturer	Positive: ≥88 ng Hb/mL. This was locally assigned by researchers and approved by the Ministry of Health.
	Were the analysts blinded (masked) to the results of the reference investigation and other clinical information?	Yes

CV, coefficients of variation; df, degrees of freedom; iFOBT, immunochemical fecal occult blood test.
* Information available from manufacturer or supplier.

Supplementeray Table 1b OC-Sensor Standard for Fecal Immunochemical Tests for hemoglobin (FITTER) checklist[35]

Topic	Item	Documentation
Specimen collection and handling		
	Name of specimen collection device and supplier (address).	Name collection device: S-bottle OC-Auto sampling bottle 3 Eiken Chemical Co LTD, 4-19-9 Taito, Taito-ku, Tokyo, 110-8408, Japan
	Description of specimen collection device (vial with probe/stick, card, other).	Flat tube with collection stick immerged in a preservative solution
	Description of specimens used if an *in vivo* study (single or pooled faeces, artificial matrix with added blood, etc).	Single human faeces sample
	Details of fecal collection method (sampling technique and number of samples).	Ribbed section of the sampling stick is striked 4 times through the stool.
	Who collected the specimens from the samples (patient, technician, etc).	Participant
	Number of fecal specimens used in the study (single, pooled, individual patient faeces).	Single sample of individual patient faeces
	Mean mass of faeces collected.*	10 mg
	Volume of buffer into which specimen is taken by probe, applicator stick or card.*	2 mL
	Time and storage conditions of fecal specimen from "passing" to sampling, including time and temperature (median and range).	Analysis took place at same day of arrival of the FIT in the lab and the FIT was kept by ambient air temperature.
	Time and storage of collection devices from specimen collection to analysis, including time and temperature (median and range). A concise description of process from collection to analysis is recommended.	Participants were asked to post the faeces samples within 24 hours after collection and keep the sample in the refrigerator. The date of sample collection is noted. FIT was transported and analyzed by ambient air temperature. All samples were analyzed within 5 days after collection.
Analysis		
	Name of analyzer, model, supplier (address), number of systems if more than one used.	Diana OC Sensor, serial number SN N00738. Supplier: Eiken Chemical Co LTD, 4-19-9 Taito, Taito-ku, Tokyo, 110-8408, Japan
	Number of times each sample was analyzed.	Single or twice. If first analysis resulted in 'no results' analysis was repeated.
	Analytical working range* and whether samples outside this range were diluted (factor) and re-assayed.	50- 1000 ng/mL. Samples were not diluted.
	Source of calibrator(s) (supplier with address), number of calibrator(s), how concentrations were assigned* and details of calibration process including frequency.	Supplier: Eiken Chemical Co LTD, 4-19-9 Taito, Taito-ku, Tokyo, 110-8408, Japan. 1 calibrator measuring 6 dilutions. Calibration was done before start of the study. The same lot number was used during the study period. Control low: range 120 and 173 ng Hb/mL, lot number 5Z017 Control high: range 379 and 513 ng Hb/mL, lot number 5Z017 If all values were within the margins, the test samples were run for analysis.

Analytical imprecision*, ideally with number of samples analyzed, concentrations, and mean, SD and CV.	In total, 20 measurements were done for each control in 5 days. Each day 2 control were run once in the morning and evening. Control low: range concentrations measured: 148-178 ng Hb/ mL, mean 159.2 ng Hb/ mL, S total measured: 9.2 ng Hb/ mL (claim S total 15), total precision CV 5.8%. S within 9.5 (claim S total 10), within run precision CV: 6.0%. Control high: range concentrations measured: 389-512 ng Hb/ mL, mean 468.3 ng Hb/ mL, S total measured: 4.9 ng Hb/ mL (claim S total 32.8), total precision CV 1.0%. S within 20.0 (claim S total 23.4), within run precision CV: 4.3%.
Quality management	
Source (address) or description of internal quality control materials, number of controls, assigned target concentrations and ranges, how target concentrations were assigned, rules used for acceptance and rejection of analytical runs.	Daily two rounds of control were done before analytical runs. Both controls should be right before start of analysis. If still not right calibration should follow (this has not happened during the study).
Participation in external quality assessment schemes: (name and address of scheme), frequency of challenges, performance attained.	Not applicable for study period.
Accreditation held by the analytical facility (address).	Accreditation by CCKL, Mariaplaats 21-D, 3511 LK UTRECHT
The number, training and expertise of the persons performing the analyses and recording the results	4 trained technician's
Result handling	
Mode of collection of data- manual recording or via automatic download to IT system, single or double reading	Data from analyzer were uploaded by usb-stick to the automatic IT system.
Units used, with conversions to µg Hb/g faeces if ng Hb/ mL used.	In analyzing and reporting results ng Hb/ mL was used. For reporting in publications this is converted to µg Hb/g faeces.[200]
Cut-off concentration(s) if used and explanation of how assigned locally or by manufacturer	Positive: ≥75 ng Hb/ mL. This was assigned by researchers and approved by the Ministry of Health.
Were the analysts blinded (masked) to the results of the reference investigation and other clinical information?	Yes

CV, coefficients of variation; df, degrees of freedom; iFOBT, immunochemical fecal occult blood test.
* Information available from manufacturer or supplier.

4

Participation and ease of use in colorectal cancer screening: a comparison of two fecal immunochemical tests

C.M. de Klerk, E. Wieten, A. van der Steen, C.R. Ramakers, E.J. Kuipers, B.E. Hansen, I. Lansdorp-Vogelaar, P.M. Bossuyt, M.C.W. Spaander, E. Dekker

American Journal of Gastroenterology 2019

ABSTRACT

Introduction

The impact of fecal immunochemical test (FIT) based colorectal cancer (CRC) screening on disease incidence and mortality is affected by participation, which might be influenced by ease of use of the FIT. We compared participation rates and ease of use of two different FITs in a CRC screening program.

Methods

Two study designs within the Dutch CRC screening program. In a paired cohort study, all invitees received two FITs (OC-Sensor, Eiken, Japan and FOB-Gold, Sentinel, Italy) and were asked to sample both from the same stool. Ease of use of both FITs was evaluated by questionnaire. In a randomized controlled trial, invitees were randomly allocated to receive one of the two FITs to compare participation and analyzability.

Results

Of 42,179 invitees in the paired cohort study, 21,078 (50%) completed two tests and 20,727 (98%) returned the questionnaire. FOB-Gold was reported significantly easier to use. More participants preferred FOB-Gold (36%) than OC-Sensor (5%), yet most had no preference (59%, $p<0.001$). In the randomized trial, 936 of 1923 invitees (48.7%) returned the FOB-Gold and 940 (48.9%) the OC-Sensor, a difference of -0.2% (CI: -3.4% to 3.0%), well within the pre-specified 5% non-inferiority margin ($p=0.001$). Only one FOB-Gold (0.1%) and four OC-Sensors (0.4%) were not analyzable ($p=0.18$).

Discussion

While FOB-Gold was significantly but marginally considered easier to use than OC-Sensor, the number of analyzable tests and participation rates in organized CRC screening are not affected when either FIT is implemented as a primary screening test.

INTRODUCTION

Population-based screening using guaiac fecal occult blood testing (gFOBT) reduces colorectal cancer (CRC)-related mortality.[1] Since several years, gFOBT is being replaced by a quantitative and more sensitive FOBT, the fecal immunochemical test (FIT). Compared to gFOBT, FIT is a one sample test, that consists of a probe and a tube instead of a smear card, and does not require dietary restrictions.[2, 3] Consequently, participation rates are higher with FIT than gFOBT.[4, 5] FIT is currently recommended in European guidelines as the test of choice for CRC population screening.[2] Multiple FITs are however available for screening, varying in the sampling tube design, buffer volume and sampling instructions.[3] Despite the clear advantages of FIT over gFOBT, specific differences between FITs could also affect participation.

So far, comparative evidence on the effects on participation for the available FITs in organized population-based screening is limited. Little higher participation rates were observed with the OC-Sensor (Eiken, Japan) than with FOB-Gold (Sentinel, Italy) in Spanish, French and Latvian screening settings (62% vs 59%; 40% vs 38%; 47% vs 45%), but a Dutch pilot screening program observed no differences (63% vs 63%).[6-9]

Some previous studies used the preferences of screenees to assess differences in ease of use, but they either compared FIT versus gFOBT[4, 10, 11], single versus multiple FIT samples [12], or were performed in a small, non-screening setting [12]. Overall, these studies confirmed a preference for a one sample test, without dietary instructions, a sampling probe and a tube for fecal collection, over a card-based test.

Ease of use of FOB-Gold and OC-Sensor, two of the most frequently used FIT brands, has so far only been evaluated by comparing the number of returned non-analyzable tests. In those studies, results were in favor of OC-Sensor.[6, 8, 9] Since then, the FOB-Gold testing tube has been modified to facilitate opening the tube and to lower the number of non-analyzable tests due to a loss of buffer.[8]

We recently showed that the accuracy in detecting advanced neoplasia (AN) is comparable for OC-Sensor and FOB-Gold.[13] To facilitate further informed decision making about the choice of FIT in population-based screening programs, additional evidence on other aspects of FIT that could influence screening effectiveness, such as ease of use and participation, is needed. At request of the Dutch minister of Health, we performed a large cohort study in which screening invitees were asked to complete both OC-sensor and FOB-Gold and assess ease of use and preference of FIT brand by questionnaire. In parallel, a randomized trial compared participation rates and the proportions of non-analyzable tests between OC-sensor and FOB-Gold.

METHODS

Study population

This study was embedded within the Dutch CRC screening program between May 2016 and March 2017. The Dutch CRC screening program started in 2014. It gradually invites 55 to 75-year old individuals for biennial FIT screening with the FOB-Gold test. Details have been described previously.[14] The target population for our study consisted of all individuals in the South-West region of The Netherlands that that were eligible for first-round screening in 2016. They were aged 59, 61, 63, 71 and 75. Exclusion criteria were identical to those in the Dutch CRC screening program: those with either a life expectancy of less than five years, a past proctocolectomy, current treatment for CRC and a history of inflammatory bowel disease, were not invited for colonoscopy at consultation after a positive FIT.[15] Organization and procedures were according to the Dutch CRC screening quality guidelines.[16]

Study design

We conducted a paired cohort study, in which invitees received two FITs (FOB-Gold and OC-Sensor) and a randomized trial, in which invitees received only one FIT, either FOB-Gold or OC-Sensor. To select invitees for both trials, random samples were taken from the same target population using a computer-generated algorithm (SPSS, IBM, version 23). Individuals selected for the randomized trial were not invited for the paired cohort study, and vice versa.

In the paired cohort study, invitees were instructed to sample each FIT from the same bowel movement. In the randomized trial, invitees received either a FOB-Gold or an OC-Sensor. All FITs were sent to the invitees with postal mail. This invitation package additionally included detailed sampling instructions for each test, as prescribed by the specific manufacturer, and an informed consent form. Invitees were asked to complete the consent and to fill in the stool sampling date.

Invitees for the paired cohort study additionally received a questionnaire about the ease of use for the two FITs (Supplementary Material 1). The questionnaire included seven factual questions focused on: 1. clarity of instructions on how to open the test, 2. ease of opening the test, 3. ease of using the stick, 4. ease of replacing the stick in the tube, 5. ease of closing the tube with the cap, 6. clarity of the sampling instructions and 7. the preferred test. Question 1-6 could be answered on a five-point scale, anchored at "1. Totally agree" to "5. Totally disagree". Question 7 invited a preference for either the "round test" (FOB-Gold), the "flat test" (OC-Sensor) or "no preference". In addition, we invited participants to clarify any reasons for their preference. A copy of the questionnaire is included in the supplementary material.

Consenting invitees were asked to return the FIT(s) and the consent form within three days after stool sampling to a specialized laboratory for analysis in a study specific sealed and pre-paid envelope. Invitees in the paired cohort study were asked to additionally return the questionnaire in the same envelope. Returned FITs and informed consent forms were checked by specialized laboratory staff. Questionnaires were scanned electronically; results were automatically uploaded to a database.

Characteristics of FOB-Gold and OC-Sensor tests

FOB-Gold and OC-Sensor have similar mechanisms for detecting blood in feces, based on antibodies to human globin. The difference in sampling tubes is illustrated in Figure 1.[3] The FOB-Gold is a round tube, containing 1.7 ml preservative buffer, with a wide opening and two screw caps at each end of the tube, one green cap to which the collection probe (stick) is attached and one transparent cap that is used for analysis in the laboratory. The probe ends in a serrated tip that should be inserted in the stool sample at four different places and replaced in the tube. The OC-Sensor is a flat tube, containing 2.0 ml preservative buffer, with a narrow opening and one green 'clicking' cap. The collection probe is attached to the cap, has a serrated tip and should be scraped through the stool sample in four different areas. After replacing the probe in the tube and closing the cap, participants are instructed to sway the tube to assure that the sample is fully suspended in the buffer.

Statistical analysis

Excluded from paired analysis were participants with incomplete tests: one or two non-analyzable tests (due to fecal overload, loss of 2/3 or more of the total buffer volume both from visual assess by laboratory staff and by automatic system), missing barcode or another technical problem) and participants with an unreliable test result (in case the return date was more than six days after sampling or if the sampling date was missing).[13] Differences between FITs in terms of ease of use were compared in the paired cohort study using the Wilcoxon signed-rank test statistic. To evaluate the existence of subgroup differences (sex, age or socioeconomic status) in the preference for either FOB-Gold, OC-Sensor, or neither, and to show the magnitude of any such differences we compared their reported FIT preference using chi-square statistics.

For each arm of the randomized trial, the participation rate was calculated by dividing the number of participants returning the FIT by the total number of invitees. We calculated an absolute and a relative difference in participation rate between FOB-Gold and OC-Sensor with corresponding 95% confidence intervals. We hypothesized there would be no substantial difference in participation rate between FOB-Gold and OC-Sensor, the test previously used in the Dutch pilot screening program. In testing this, we used a 5% non-inferiority margin. Proportions of non-analyzable tests were also assessed in the

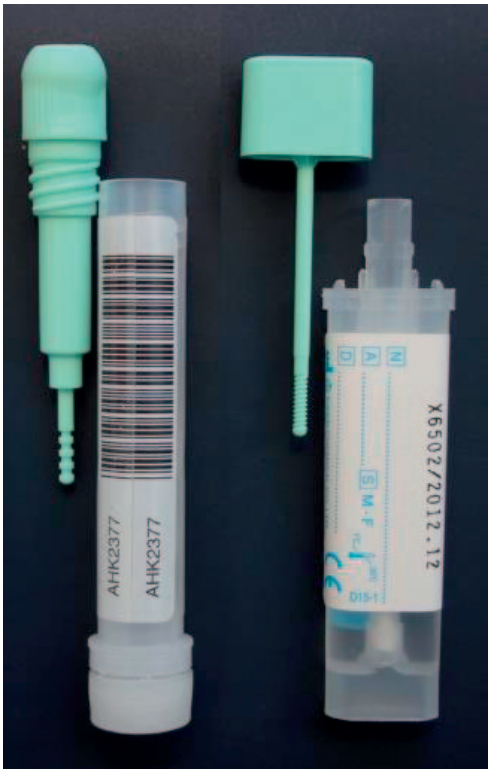
randomized trial and compared using a chi-square test. Non-analyzability could be due to an unreadable barcode, a broken tube, missing buffer, a too large sample or too small sample, a missing sample or another reason that made analysis technically impossible.

For subgroup analyses, participants were categorized into two age groups: 55-64 years and 65-75 years, because the selected screening invitees in 2016 consisted of five specific birth years. Socioeconomic status was assessed by the area social status score (combining education, income and employment status) developed by the Netherlands Institute of Social Research[17], and grouped into quintiles, with 1 being the highest status and 5 being the lowest. In all tests except the non-inferiority test for participation, two-sided p-values of 0.05 were considered significant. For participation, a one-sided 0.025 significance level was used.

Sample size calculations

The sample size for the paired trial was guided by the aim to evaluate differences in diagnostic yield, relative sensitivity, and relative specificity, as reported in detail elsewhere.[13]

Figure 1 FOB-Gold (left) and OC-Sensor (right) fecal immunochemical test



Modified from Schreuders et al. 2016 [3]

In the randomized trial, our objective was to yield estimates of the relative participation rate of both FITs in CRC screening with sufficient precision to allow national decision making on purchasing and implementing one of both tests. We assumed that 50% of invitees would participate in the randomized trial. Assuming no difference in participation between the two FITs, 2019 individuals were required in each arm of the randomized trial to have 90% power in excluding an absolute difference of 5% or more in participation, using a significance level of 0.025.

Ethical approval

The study protocol was evaluated by the Dutch National Health Council and approved by the Minister of Health (Population Screening Act; no. 769500-135716-PG; date 4 June 2015) and registered before its initiation at the Dutch Trial Registry as trail no. NTR5874.

RESULTS

Study participants

The study flow is summarized in Figure 2. Baseline characteristics of participants in the paired cohort study and the randomized trial are described in Table 1. Half of the participants were male and median age was 60 (IQR 59-61). There were significant differences between participants in the randomized trial and the paired cohort. The latter consisted of more elderly participants, and more participants with a higher socioeconomic status ($p<0.001$). There were no differences in between the two arms of the randomized trial.

Ease of use

Of 42,179 individuals invited to the paired cohort study, 21,078 participated and returned two completed FITs, of which 20,727 (98%) also returned the questionnaire (Figure 2). Reported answers on ease of use are shown in Table 2 and Figure 3.

Significant but small differences were found for almost all aspects of ease of use, in favor of FOB-Gold, except for the clarity of instructions for opening the test ($p=0.34$). The largest difference was found for replacing the stick in the tube, with 94% of participants 'totally agreeing' or 'agreeing' this was easy with FOB-Gold versus 79% indicating the same for OC-Sensor, resulting in an absolute difference of 15%. A difference of 5% was found for ease of closing the test with the cap: 98% said this was easy ('totally agree' and 'agree') for FOB-Gold compared to 93% for OC-Sensor. Smaller and sometimes tiny differences were observed for the other domains: ease of using the stick, ease of opening the test and clarity of the sampling instructions.

Figure 2 Flow of participants in the paired cohort study and in the randomized controlled trial.

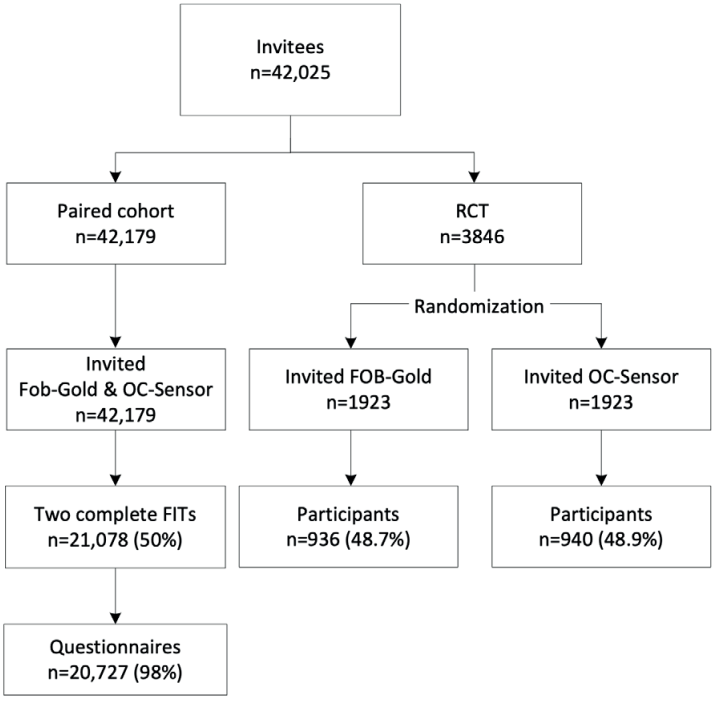


Table 1 Baseline characteristics per study arm

	Paired cohort	p-value**	RCT	
	Total n=20,727		FOB-Gold n=936	OC-Sensor n=940
Male sex, n (%)	10,425 (50%)	1.0	475 (51%)	472 (50%)
Female sex, n (%)	10,302 (50%)		461 (49%)	468 (50%)
Age in years median (IQR)*	60 (58-62)		60 (59-61)	60 (59-61)
55-64, n (%)	17,031 (82%)	<0.001	848 (91%)	841 (89%)
65-75, n (%)	3696 (18%)		88 (9%)	99 (11%)
SES, n (%)		<0.001		
Very high	4152 (20%)		46 (5%)	49 (5%)
High	4570 (22%)		117 (13%)	122 (13%)
Average	3819 (19%)		209 (22%)	217 (23%)
Low	4845 (23%)		376 (40%)	374 (40%)
Very low	3294 (16%)		181 (19%)	176 (19%)
Missing	47 (<1%)		7 (1%)	2 (<1%)

*Age at time of fecal immunochemical test (FIT) invitation

** Difference between study group in RCT and paired cohort

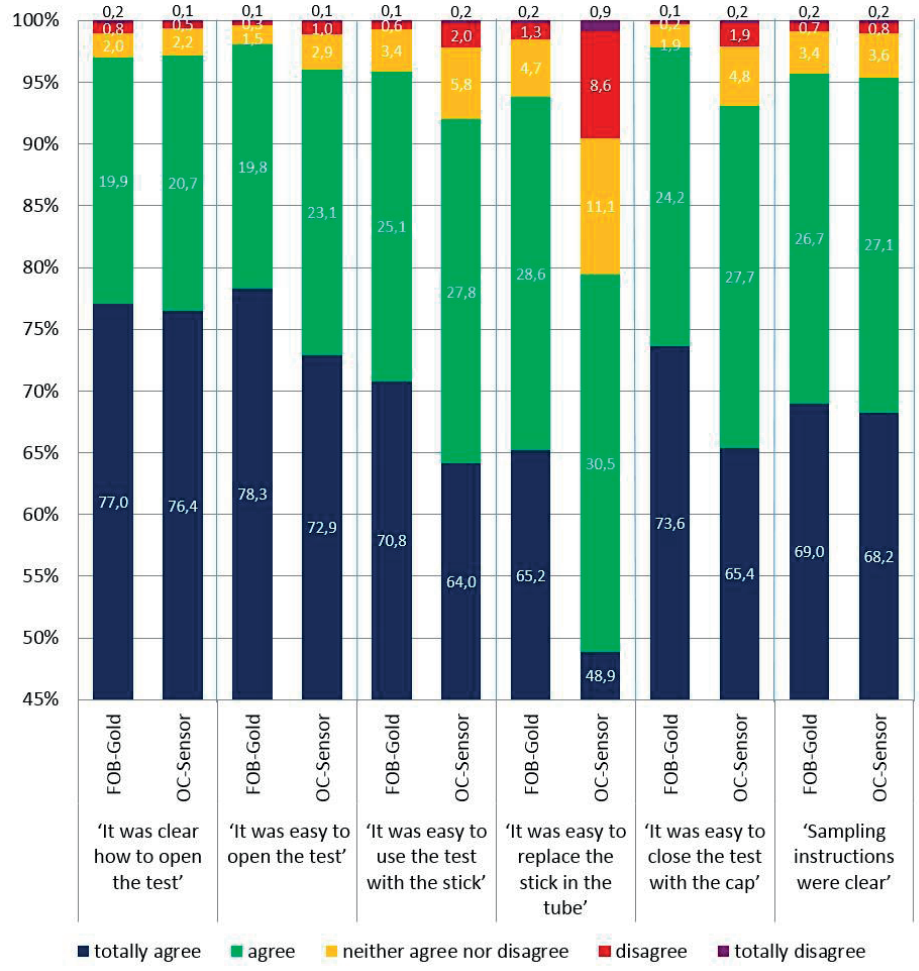
IQR: interquartile range; SES: socioeconomic status; RCT: randomised controlled trial

Table 2 Responses per aspect of ease of use for FOB-Gold and OC-Sensor

	FOB-Gold	OC-Sensor	p-value*
'It was clear how to open the test'	n=17,821	n=17,710	0.34
Totally agree	13,731 (77.0%)	13,524 (76.4%)	
Agree	3552 (19.9%)	3673 (20.7%)	
Neutral	362 (2.0%)	395 (2.2%)	
Disagree	146 (0.8%)	92 (0.5%)	
Totally disagree	40 (0.2%)	26 (0.1%)	
'It was easy to open the test'	n=17,794	n=17,688	<0.001
Totally agree	13,927 (78.3%)	12,896 (72.9%)	
Agree	3518 (19.8%)	4086 (23.1%)	
Neutral	275 (1.5%)	509 (2.9%)	
Disagree	55 (0.3%)	171 (1.0%)	
Totally disagree	19 (0.1%)	26 (0.1%)	
'It was easy to use the test with the stick'	n=17,760	n=17,644	<0.001
Totally agree	12,567 (70.8%)	11,298 (64.0%)	
Agree	4460 (25.1%)	4912 (27.8%)	
Neutral	600 (3.4%)	1,032 (5.8%)	
Disagree	110 (0.6%)	359 (2.0%)	
Totally disagree	23 (0.1%)	43 (0.2%)	
'It was easy to replace the stick in the tube'	n=17,783	n=17,681	<0.001
Totally agree	11,587 (65.2%)	8648 (48.9%)	
Agree	5084 (28.6%)	5392 (30.5%)	
Neutral	843 (4.7%)	1957 (11.1%)	
Disagree	236 (1.3%)	1520 (8.6%)	
Totally disagree	33 (0.2%)	164 (0.9%)	
'It was easy to close the test with the cap'	n=17,787	n=17,684	<0.001
Totally agree	13,095 (73.6%)	11,563 (65.4%)	
Agree	4304 (24.2%)	4907 (27.7%)	
Neutral	334 (1.9%)	841 (4.8%)	
Disagree	42 (0.2%)	330 (1.9%)	
Totally disagree	12 (0.1%)	43 (0.2%)	
'Sampling instructions were clear'	n=17,755	n=17,652	<0.001
Totally agree	12,248 (69.0%)	12,036 (68.2%)	
Agree	4749 (26.7%)	4787 (27.1%)	
Neutral	607 (3.4%)	641 (3.6%)	
Disagree	120 (0.7%)	148 (0.8%)	
Totally disagree	31 (0.2%)	40 (0.2%)	

* p-values based on Chi-square test statistics

Figure 3 Responses per aspect of ease of use for FOB-Gold and OC-Sensor



Preferred FIT

Most participants (59%) did not have a clear preference for either FIT brand, of those with a preference, more participants preferred FOB-Gold (36%) than OC-Sensor (5%) (p <0.001%; Table 3). Males preferred FOB-Gold slightly more often than females (37% versus 35%) and females preferred OC-Sensor more often (5.7% versus 4.5%) (p<0.001). Participants over 65 years of age were more frequently indifferent compared to younger participants (67% versus 57%; p<0.001), as were participants with a lower socioeconomic status, compared to those with a higher status (p<0.001).

Table 3 Preferred FIT brand as reported by participants, stratified by sex, age and socioeconomic status

	Total n=17,333	FOB-Gold n=6268 (36%)	OC-Sensor n=881 (5%)	No preference n=10,184 (59%)	p-value
Sex n=17,333					<0.001
- Males	8785	3244 (37%)	391 (4%)	5150 (59%)	
- Females	8548	3024 (35%)	490 (6%)	5043 (59%)	
Age* n=17,333					<0.001
- 55-64	14,365	5408 (38%)	761 (5%)	8196 (57%)	
- 65-75	2968	860 (29%)	120 (4%)	1988(67%)	
SES n= 17,295					
- Very high	3556	1345 (38%)	210 (6%)	2001 (56%)	<0.001
- High	3862	1466 (38%)	182 (5%)	2214 (57%)	
- Average	3225	1167 (36%)	166 (5%)	1892 (59%)	
- Low	3990	1364 (34%)	202 (5%)	2424 (61%)	
- Very low	2662	912 (34%)	120 (5%)	1630 (61%)	

FIT: fecal immunochemical test; SES: socioeconomic status
*Age at FIT invitation

Of the 6268 participants who preferred FOB-Gold, 1086 (17%) reported a reason for their preference. The most frequently reported reason was the wider opening of the tube, which made it easier to replace the sampling probe. In addition, the screw-cap was considered easier to open and easier and close. Taking a sample by sticking the probe in the stool, as with FOB-Gold, was reported easier than scraping, as prescribed for OC-Sensor. Moreover, some respondents indicated that less feces was sticking to the probe, making it easier to sample the right volume. The grip of the FOB-Gold was also frequently appreciated.

Of the 881 participants that preferred OC-Sensor 175 (20%) provided a reason for their preference. Some aspects of preference for OC-Sensor were similar to those reported for FOB-Gold, although in smaller proportions, for example the grip and easy closing of the cap. Especially appreciated in OC-Sensor was its single cap that avoids any confusion on which cap to open and prevents the loss of buffer. Its flat shape was also preferred to prevent the test from rolling away before and after sampling.

Participation

In the randomized trial, 936 out of 1923 invitees allocated to receive an FOB-Gold participated (48.7%) versus 940 out of 1923 (48.9%) allocated to OC-Sensor (RR = 1.00; 95% CI: 0.93 to 1.06) (Figure 2). The absolute difference was -0.2% (95% CI: -3.4% to 3.0%). The null hypothesis of a difference of 5% or more was rejected, demonstrating that FOB-Gold was non-inferior to OC-Sensor in terms of participation rate (p=0.001).

Non-analyzability

One out of 936 FOB-Gold tests (0.1%) could not be analyzed versus four out of 940 OC-Sensor tests (0.4%; p= 0.18).

DISCUSSION

This study found small, but statistically significant, differences in ease of use in favor of FOB-Gold compared with OC-Sensor. FOB-Gold was more often preferred than OC-Sensor, but most participants did not express a clear preference for either FIT. Despite these differences, our randomized trial showed that participation in the Dutch population based CRC screening program was not influenced by the type of FIT offered with the invitation sent by postal mail.

This study was conducted within the logistics of the Dutch CRC screening program and included a large and representative sample of the screening population. Because the study population was screening-naive, their expressed preferences were not influenced by a previous experience with one of the tests. With the large study group, we had enough power to detect small differences in preferred aspects of use. Nevertheless, some limitations have also to be acknowledged. Due to a limited number of screening-naive individuals in our study population in 2016, fewer were invited in each arm of the randomized controlled trial than we anticipated in our sample size calculations (1923 instead of 2019). Despite this failure to reach the targeted number, we could reject inferiority of FOB-Gold in terms of participation. Ease of use and preferences were evaluated in a large study with paired design and the effects on participation in a randomized trial, in which invitees were randomly allocated to one of both tests. These were two different groups, but both groups were randomly selected from the same target screening population, facilitating generalizability. Participants in our paired cohort, in which ease of use was assessed, were somewhat older and more had a higher socioeconomic status than in the randomized trial. Since the two arms in the randomized trial were balanced, we feel confident to conclude that any of the differences in ease of use have not affected participation. Our questionnaire had been tested in a previous study [18], but was not validated for this comparative evaluation. We do not expect any selection bias in the responding participants because the participation rate in the paired cohort was similar to that in the RCT, and 98% of participants also returned the questionnaire. The specific aim of this study was to compare FOB-Gold and OC-Sensor because these FITs were implemented in the Dutch CRC screening programme. Although these FIT brands are among all available FITs widely used ones, our results cannot be generalized to possible differences between other FIT-brands. Although our screening naive study population had no experience with one of the FITs, we cannot exclude that participant's responses might have been influenced by the fact that FOB-Gold is the test that is currently used in the national Dutch screening program.

The ease of use was appreciated differently for each FIT for almost all aspects, favoring FOB-Gold. Most differences were small and might be considered clinically irrelevant, though the higher reported ease of replacing the sampled probe into the FOB-Gold tube was evident. Earlier studies comparing FOB-Gold and OC-Sensor, did not rely on a survey with paired design to assess ease of use, but instead evaluated the number of non-analyzable tests and found higher error rates for FOB-Gold.[6, 8, 9] We found very few non-analyzable tests, and no difference in proportions by FIT. The main reasons for non-analyzability of FOB-Gold in former studies were fecal overload[8], wrong opening of the tube and loss of buffer [9] or 'incorrect sample manipulation'[6]. In our study the only non-analyzable FOB-Gold tube was missing its buffer, despite adaptations of the FOB-Gold's caps, designed to avoid opening of the wrong side of the tube.

Combining analyzability and appreciated ease of use, suggestions for optimal test design can be derived. Loss of buffer with FOB-Gold is probably due to the presence of two caps, one at each end of tube; this was reported to be more impractical by some of the participants. The wide opening of the FOB-Gold tube is appreciated but may lead, in rare cases, to over-sampling. On the other hand, a small opening in OC-Sensor can lead to under-sampling and appears to be less well appreciated in our study. Considering the other reported evaluations of participants on FIT shape, sampling instruction, cap opening and closure of these two FITs, the ultimate design could be envisioned. An ideal FIT would then be square shaped, have a wide opening, one screw cap, and sampling is instructed by sticking the probe in the stool. Possibly, the ideal FIT would positively affect the proportion of analyzable FITs.

It is known that participation in CRC screening is generally lower for persons with lower socioeconomic status and for ethnic minorities, while CRC incidence and mortality are higher in these groups.[19-21] To avoid any increase in health inequities within the population through screening, any difference in FITs that influences participation by socioeconomic or ethnic subgroup is of relevance. Our results show differences in preferences between subgroups by sex, age and socioeconomic status, but these were small and inconsequential for the screening uptake rate. Hence, we would currently not endorse using different tests for different sexes, ages or socioeconomic status. Our results are however based on the preferences of those participating in screening and did not include information on participant's ethnic background. Evidence on culturally, ethnically or racially related barriers to engage in FOBT- screening is limited, but a disgust of stool handling was included in the top five barriers in an ethnically and racial diverse population.[22] Although future study initiatives that specifically address culturally or ethnically derived barriers to engage in FIT screening could be valuable to optimize CRC screening participation, we do however not expect big differences in preference between FITs for different ethnic groups.

The absence of a difference in participation rate for OC-Sensor or FOB-Gold within an organized CRC screening program confirms the results of a previous study in the pilot program in the Netherlands, in which the uptake was 63% with both tests.[8] At time of the pilot program, no alternative screening option existed in the Netherlands, while invitees in our study had the option to participate in the national screening without participating in this study. This probably explains a slightly lower participation rate in our evaluation compared with the former study. Other studies have shown significant though small differences in favor of OC-Sensor: 59% compared with 62%[6] and 45% versus 47%.[9] Whether participation rates in successive screening rounds is influenced by the FIT that was distributed in the first round is not yet known, but may be relevant for the evaluation of ongoing screening programs.

Continuous evaluation of new strategies to optimize screening program effectiveness and reduce possible harms is crucial to guarantee screening quality, and should be performed on a routine basis.[23] The absence of a difference in participation rates between FOB-Gold and OC-Sensor in this study, despite a preference for one particular test in a minority of participants and differences in reported ease of use, shows that these small differences do not affect the willingness to engage in CRC screening. This result, in combination with the evidence that both FITs have an equivalent accuracy to detect AN [13], implies that other considerations can guide the selection of a test in population-based screening programs for CRC. Cost-effectiveness analyses, efficiency in logistics and efficiency in laboratory analyses are other considerations that should be taken into account.

Acknowledgements

The authors would like to thank the Netherlands Organisation for Health Research and Development of the Dutch Ministry of Health (ZonMw) for funding, and all participants for their efforts to perform two tests and a questionnaire. The authors thank all involved coworkers of the Foundation of Population Screening South-West (Bevolkingsonderzoek Zuid-West), and all involved co-workers at Daklapack packaging for their contributions to the study. Finally, the authors thank Star lab (Rotterdam, The Netherlands), for their efforts in analyzing twice as many FITs compared to the nationwide Dutch screening program during the study period.

REFERENCES

1. 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.
2. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
3. Schreuders EH, Grobbee EJ, Spaander MCW, Kuipers EJ. Advances in Fecal Tests for Colorectal Cancer Screening. Current Treatment Options in Gastroenterology. 2016;14(1):152-62.
4. Deutekom M, van Rossum LG, van Rijn AF, Laheij RJ, Fockens P, Bossuyt PM, et al. Comparison of guaiac and immunological fecal occult blood tests in colorectal cancer screening: the patient perspective. Scand J Gastroenterol. 2010;45(11):1345-9.
5. Moss S, Mathews C, Day TJ, Smith S, Seaman HE, Snowball J, et al. Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England. Gut. 2017;66(9):1631-44.
6. Zubero MB, Arana-Arri E, Pijoan JI, Portillo I, Idigoras I, Lopez-Urrutia A, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test. Front Pharmacol. 2014;4:175.
7. Faivre J, Dancourt V, Denis B, Dorval E, Piette C, Perrin P, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer. 2012;48(16):2969-76.
8. 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. 2016;66(11):1975-82.
9. Santare D, Kojalo I, Huttunen T, Rikacovs S, Rucevskis P, Boka V, et al. Improving uptake of screening for colorectal cancer: a study on invitation strategies and different test kit use. Eur J Gastroenterol Hepatol. 2015;27(5):536-43.
10. Cole SR, Young GP, Esterman A, Cadd B, Morcom J. A randomised trial of the impact of new faecal haemoglobin test technologies on population participation in screening for colorectal cancer. J Med Screen. 2003;10(3):117-22.
11. Ellis RJ, Wilson S, Holder RL, McManus RJ. Different faecal sampling methods alter the acceptability of faecal occult blood testing: a cross sectional community survey. Eur J Cancer. 2007;43(9):1437-44.
12. Pham R, Cross S, Fernandez B, Corson K, Dillon K, Yackley C, et al. "Finding the Right FIT": Rural Patient Preferences for Fecal Immunochemical Test (FIT) Characteristics. J Am Board Fam Med. 2017;30(5):632-44.
13. 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.
14. 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. 2017;152(4):767-75.e2.
15. National Institute for Public Health and the Environment (RIVM). Exclusion criteria colonoscopy (Dutch)http://www.rivm.nl/Documenten_en_publicaties/Professioneel_Praktisch/Richtlijnen/Preventie_Ziekte_Zorg/Darmkanker/Exclusiecriteria_coloscopie (date accessed 20th June 2017); 2013.

16. National Institute for Public Health and the Environment (RIVM). Uitvoeringskader Bevolkingsonderzoek Darmkanker (Dutch). www.bevolkingsonderzoekdarmkanker.nl/downloads (date accessed 20th June 2017); 2017.
17. 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.
18. 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.
19. de Klerk CM, Gupta S, Dekker E, Essink-Bot ML. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. Gut. 2017.
20. Doubeni CA, Laiyemo AO, Major JM, Schootman M, Lian M, Park Y, et al. Socioeconomic status and the risk of colorectal cancer: an analysis of more than a half million adults in the National Institutes of Health-AARP Diet and Health Study. Cancer. 2012;118(14):3636-44.
21. Williams R, White P, Nieto J, Vieira D, Francois F, Hamilton F. Colorectal Cancer in African Americans: An Update. Clin Transl Gastroenterol. 2016;7(7):e185.
22. Jones RM, Woolf SH, Cunningham TD, Johnson RE, Krist AH, Rothemich SF, et al. The relative importance of patient-reported barriers to colorectal cancer screening. Am J Prev Med. 2010;38(5):499-507.
23. Bell KJ, Bossuyt P, Glasziou P, Irwig L. Assessment of changes to screening programmes: why randomisation is important. BMJ. 2015;350:h1566.

Supplementary Material 1 Questionnaire

Vragen testbuisjes.

Wij vragen u voor beide testbuisjes onderstaande vragen in te vullen.

	Rond testbuisje	Plat testbuisje
1. Het was duidelijk hoe het testbuisje te openen	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens
2. Het was makkelijk om het testbuisje te openen	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens
3. Het was makkelijk om de test uit te voeren met het staafje	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens
4. Het was makkelijk om het staafje met de ontlasting in het testbuisje te doen	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens
5. Het was makkelijk om het dopje weer op het testbuisje te doen	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens
6. Het instructieboekje bij de test was duidelijk	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens	☐ helemaal mee eens ☐ mee eens ☐ neutraal ☐ mee oneens ☐ helemaal mee oneens

7. Welk testbuisje heeft uw voorkeur om te gebruiken?

☐ Rond testbuisje
☐ Plat testbuisje
☐ Geen voorkeur

5

Performace of two faecal immunochemical tests in detecting advanced neoplasia at different positivity thresholds: a cross sectional study of the Dutch national colorectal cancer screening programe

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

Lancet Gastroenterology & Hepatology 2018

ABSTRACT

Background

Faecal immunochemical tests (FITs) are recommended for colorectal cancer screening. Two frequently used FIT methods (FOB-Gold, Sentinel Diagnostics, Milan, Italy and OC-Sensor, Eiken Chemical, Tokyo, Japan) perform similarly in detecting advanced neoplasia (ie, colorectal cancer and advanced adenoma) at a fixed positivity cutoff for haemoglobin concentration. It is unclear whether the performance of the two methods is comparable at other thresholds. We compared the accuracy of the two assays in detecting advanced neoplasia across various thresholds.

Methods

In a cross-sectional study in the Dutch national screening programme, individuals who were screening naïve (age 55-75 years) living in the southwest region of the Netherlands were invited to use two different FIT assays on the same bowel movement. Eligible participants were randomly selected from municipal registers. Participants were referred for colonoscopy if either FIT assay result met the predefined positivity threshold (≥15 µg haemoglobin per g faeces). We compared the respective distributions of reported hemoglobin concentration and positivity rates with various FIT positivity thresholds. The performance of each FIT for identifying advanced neoplasia at colonoscopy in FIT-positive assays was compared with the area under the receiver operating characteristic curve.

Results

21,078 (50.0%) of 42,179 invitees completed both FIT assays. The distribution of haemoglobin concentrations differed significantly between the two FITs ($p<0.0001$), with higher positivity rates for OC-Sensor at FIT thresholds of 5 and 10 µg haemoglobin per g faeces, similar positivity rates at 15 and 20 µg haemoglobin per g faeces, and higher rates for FOB-Gold at FIT thresholds of 25-150 µg haemoglobin per g faeces. 2046 (9.7%) of 21,078 participants had at least one FIT assay that was positive and of these, 1724 (84.3%) attended colonoscopy. The accuracy of results in individuals undergoing colonoscopy did not significantly differ between the FITs, with an area under the receiver operating characteristic curve of 0.675 (95% CI: 0.649 to 0.702) for FOB-Gold and 0.686 (95% CI: 0.661 to 0.712) for OC-Sensor ($p=0.40$). At identical positivity rates, the positive predictive value of the two FIT assays was similar (difference varying from 0.5% [95% CI -2.6 to 3.7] at a positivity rate of 3.5% to 2.4% [-2.5 to 7.3] at a positivity rate of 2.0%).

Conclusion

The two widely used FITs have significantly different distributions of reported hemoglobin concentration and yield different positivity rates at equal thresholds. However, they perform similarly in detecting advanced neoplasia at a preset positivity rate. When implementing

either FIT in a screening programme, the desired positivity rate that identifies participants to be referred for colonoscopy should first be set, guided by available resources and feasibility.

Funding
The Netherlands Organization for Health Research and Development (ZonMw, nr 200350001).

INTRODUCTION

Faecal immunochemical tests (FITs) are the recommended non-invasive test of choice for population-based colorectal cancer screening in Europe.[1, 2] Compared with the guaiac faecal occult blood test, FITs have a higher sensitivity in detecting colorectal cancer and advanced adenoma (ie, advanced neoplasia), especially at low haemoglobin concentrations, and provides a quantitative measurement of the faecal haemoglobin concentration.[2, 3] The quantitative measurement enables screening programmes to choose a positivity threshold that balances a high diagnostic yield (minimising the number of false-negative FIT results) while limiting the proportion of negative colonoscopies (minimising the number of false-positive FIT results), taking resource constraints, capacity, and costs into account.[4]

At least four quantitative FIT assays are now available for colorectal cancer screening and evidence for their diagnostic performance is increasing.[5] A large paired comparative study showed that two of the most widely used FITs, FOB Gold (Sentinel Diagnostics, Milan, Italy) and OC-Sensor (Eiken Chemical, Tokyo, Japan), had a similar yield in detecting advanced neoplasia at a fixed threshold of 15µg hemoglobin per g faeces.[6] It remains unknown whether this equivalence in yield extends to other positivity thresholds. This information could be useful to enable the appropriate FIT threshold to be adopted in a screening programme that uses a different positivity threshold than previously studied, or in screening programmes in which a different threshold is considered.

In a previous trial (n=12 054), in which participants were randomly allocated to either a FOB-Gold or an OC-Sensor test, different positivity rates for the two tests were found at a threshold of 10 µg haemoglobin per g faeces. However, positive predictive values for the detection of advanced neoplasia were similar when the FITs were compared at identical positivity rates.[7] We aimed to evaluate this finding in a substantially larger study with a paired design, and to explain the probable reasons for differences in positivity rates. We compared the full haemoglobin distributions of the two FIT assays, the positivity rates at equal thresholds, the performance in detecting advanced neoplasia, and the positive predictive values at identical positivity rates.

METHODS

Study population and design

This large, cross-sectional, cohort study was done between May 10, 2016 and March 1, 2017 within the Dutch national colorectal cancer screening programme. Our study was designed to compare the diagnostic yield of advanced neoplasia in two FIT assays at a fixed faecal haemoglobin concentration threshold. The structure of the Dutch national colorectal cancer screening programme and the design of this study have previously been

described in detail.[6, 8] In short, a random selection was made of screening naive individuals, eligible for screening (aged 55 to 75 years) in 2016 and living in the southwest region of the Netherlands including the provinces of Zuid-Holland and Zeeland. Individuals were eligible based on their home address and age. The random selection was made with use of a computer run algorithm (SPSS version 23). All selected individuals were sent an invitation, a consent form and two FIT assays (FOB-Gold and OC-Sensor) by post. Invitees were asked to provide a sample for both FIT assays in the same bowel movement. Participants sent the used FITs and the consent form, including sampling date, to one accredited centralized laboratory (Starlab-MDC, Rotterdam,

Netherlands) in a sealed pre-paid envelope. The study was approved by the Dutch National Health Council (Population Screening Act; publication no. 2015/09) and registered in the Dutch Trial Registry (NTR5874). All study participants gave written informed consent.

Procedures

The FOB-Gold tests were measured using a Bio Majesty JCA-BM6010/C analyser and the OC-Sensor tests by the OC-Sensor Diana analyser. Details of the analysis have been previously described.[6] Quantitative results for both tests were provided in ng haemoglobin per mL buffer and converted into µg haemoglobin per g of faeces for the purpose of this study.[9] All OC-Sensor FIT concentrations greater than 200 µg haemoglobin per g faeces were reported as more than 200 µg haemoglobin per g faeces, but FOB-Gold FIT concentrations were reported without an upper limit.

Participants were informed about their FIT results (negative or positive) by post and were invited for a precolonoscopy interview if either of both of the FIT concentrations exceeded the preset positivity threshold of 15 µg haemoglobin per g faeces. During the precolonoscopy interview, participants were excluded from colonoscopy if they had a life expectancy of 5 years or less, a proctocolectomy in the past, were under current treatment for colorectal cancer, had a history of inflammatory bowel disease, or had undergone a complete colonoscopy in the past 5 years.[10]

Colonoscopies were done in one of the population screening certified colonoscopy centers by accredited endoscopists who do at least 200 colonoscopies a year with an adenoma detection rate of 30% or more. All endoscopically identified lesions were reported using an automated structured colonoscopy reporting system. Removed lesions were sent for pathology review in separate containers.[11] Adenomas of 10 mm or more, with 25% or more villous component, or high-grade dysplasia, or a combination, were defined as advanced adenoma. Advanced neoplasia was defined as advanced adenoma or colorectal cancer.

Each participant was classified according to the most advanced lesion detected. Logistics were done in accordance with the Dutch colorectal cancer screening quality guidelines.[8]

Statistical analysis

Socioeconomic status was assessed with the Dutch area social status score and was grouped into quintiles. These scores are a composite measure, including education, income and employment status developed by the Netherlands Institute of Social Research.[12] Participation rate was calculated as the number of participants who returned at least one FIT relative to the number of invitees. Not included in this paired analysis were participants who returned one or two FITs that were not analysable (due to faecal overload, buffer loss, missing barcode or another technical problem), who had one or two unreliable test results (if the return date was more than 6 days after sampling or if the sampling date was missing), or participants in whom one test was missing.

The reported faecal haemoglobin concentration of each FIT assay was reflected in a cumulative distribution function. We evaluated the difference between the two curves with a two-sample Kolmogorov-Smirnov test statistic, using a permutation test to calculate the p value while accounting for the paired nature of the data.[13]

To further evaluate the positivity rates of each FIT at different thresholds, the reported haemoglobin concentrations were examined with use of potential thresholds with 5 µg haemoglobin per g faeces increments (5, 10, 15, and so on, up to 150 µg haemoglobin per g faeces). Positivity rates were calculated by dividing the number of positive tests by the total number of participants who completed two tests. Absolute differences in positivity rate between FOB-Gold and OC-Sensor were calculated with corresponding 95% CIs using a Wald interval with Bonett-Price adjustment.[14]

The performance of the two FITs in detecting advanced neoplasia was evaluated in all participants who attended colonoscopy after one or two positive FIT results by calculating the area under the under the receiver operating characteristic curve (AUC). We compared the AUC using the method of DeLong and collegues.[15]

The positive predictive value was defined as the proportion of participants with a positive FIT result in whom advanced neoplasia was identified at colonoscopy. Absolute differences in positive predictive values between FOB-Gold and OC-Sensor were also compared at equal positivity rates (2-6%); 95% CIs were calculated with use of a Wald interval with Bonett-Price adjustment. Calculated p values were two-sided and differences were considered significant if p values were less than 0.05. The study sample size was based on the comparison of the diagnostic yield for the two FITs, the initial objective of the study.[6]

SPSS (version 23) was used for statistical analyses, except for the permutation test, which was run in R (version 3.5.1).

Role of funding source

The Netherlands Organization for Health Research and Development of the Dutch Ministry of Health (ZonMw) funded this study but was not involved in the study design; in the data collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. CMdK, EW, IL-V, PMB and ED had access to the raw data and ED had final responsibility for the decision to submit for publication.

RESULTS

Study participants

Of 42,179 invitees, 22,064 (52.3%) participated in the study and 21,078 (50.0%) completed both FITs (Figure 1). The 986 (4.5%) participants with incomplete FIT results had either returned only one test (n=702), one or two non-analysable tests (n=54), had one or two unreliable test results (n=260); some participants also had a combination of these (n=30), and were excluded from the paired analyses. The proportion of participants with complete tests was equal between women and men (10,489 [49.8%] vs 10,589 [50.2%]). The median age of participants was 60 years (IQR 58 to 62) and socioeconomic status was classified as very low in 15.9%, low in 23.3%, average in 18.4%, high in 22.1%, very high in 20.0% and missing in 0.2% of the participants.

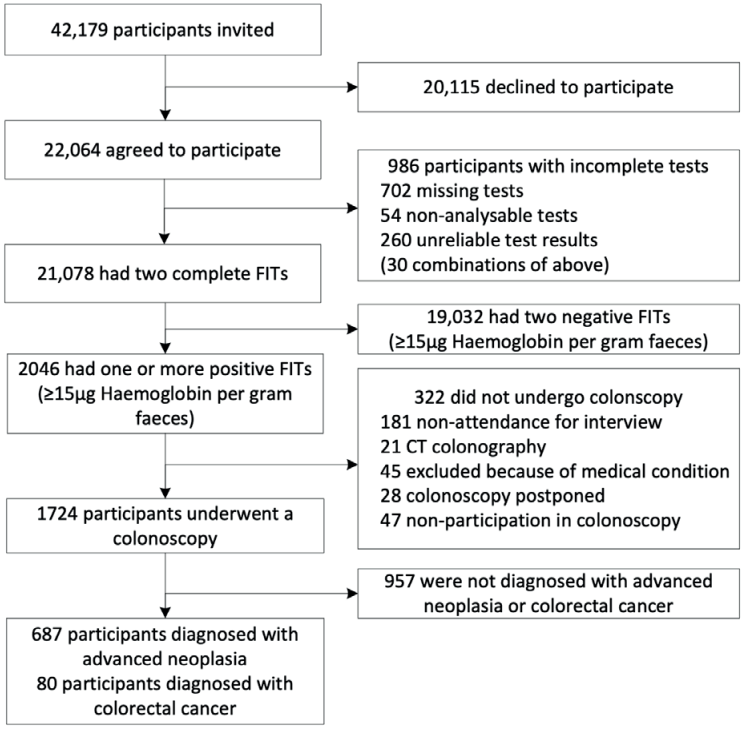
Distributions of Hb concentrations

Figure 2 shows the cumulative distributions of reported haemoglobin concentration for FOB-Gold and OC-Sensor, including the FIT results of all 21,078 participants that completed two tests. The two distributions differed significantly ($p<0.0001$), with the difference most marked at lower concentrations. In 18,438 (87.5%) of the FOB-Gold tests, no haemoglobin was detected, resulting in a median haemoglobin concentration of 0.0 µg haemoglobin per g faeces (IQR 0.0 to 0.0). With the OC-sensor test, 8070 (38.3%) test results reporting 0.0 µg haemoglobin per g faeces with a median haemoglobin concentration of 0.5 µg haemoglobin per g faeces (IQR 0.0 to 2.2). The maximum reported haemoglobin concentration with FOB-Gold was 285 µg haemoglobin per g faeces. OC-Sensor test results were capped at 200 µg haemoglobin per g faeces; 282 participants (1.3%) had an OC-Sensor test result exceeding this cap.

Positivity rates of both FIT at potential thresholds

As a direct consequence of the difference in haemoglobin distribution, positivity rates at each threshold was different for FOB-Gold and OC-Sensor (Figure 3, Supplementary Table 1). The positivity rate of FOB-Gold was 4.8% lower than that of OC-Sensor at a haemoglobin

Figure 1 Study profile



concentration threshold of 5 µg haemoglobin per g faeces (95%CI: -5.2% to -4.4%) and 0.9% lower at threshold of 10 µg haemoglobin per g faeces (95%CI: -1.2% to -0.6%). Similar positivity rates were found at thresholds of 15 µg haemoglobin per g faeces (difference -0.3%; 95%CI: -0.5% to 0.0%) and 20 µg haemoglobin per g faeces (0.2%; 95%CI: -0.1% to 0.4%). A higher positivity rate of FOB-Gold than OC-Sensor was found at thresholds 25 µg haemoglobin per g faeces (0.4%; 95%CI: 0.2% to 0.7%) up to 150 µg haemoglobin per g faeces (0.8%; 95%CI: 0.6% to 0.9%; Figure 3).

Performance of both FIT in detecting advanced neoplasia at equal positivity rates

2046 (9.7%) of 21,078 participants had at least one positive FIT, of whom 1724 (84.3%) attended colonoscopy. Discrimination in participants with one or two true-positive FIT results (at 15 µg haemoglobin per g faeces) was not significantly different for FOB-Gold and OC-Sensor with an AUC of 0.675 (95%CI: 0.649 to 0.702) for FOB-Gold and 0.686 (95%CI: 0.661 to 0.712) for OC-Sensor (p=0.40; Figure 4). This finding is reflected in the similar positive predictive value to detect advanced neoplasia at equal positivity rates for both FIT assays (Figure 5, Supplementary Table 2). The difference in positive predictive value between FOB-Gold and OC-Sensor varied from 0.5% (95%CI: -2.6% to 3.7%) at a

Figure 2 Cumulative distributions of reported haemoglobin concentrations with FOB-Gold and OC-Sensor

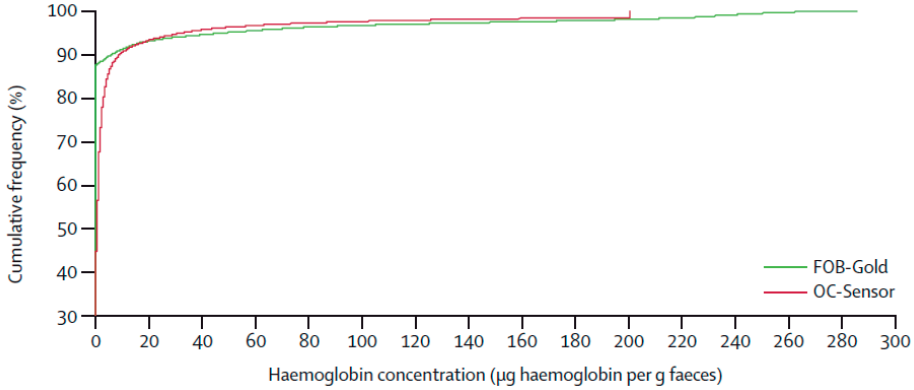
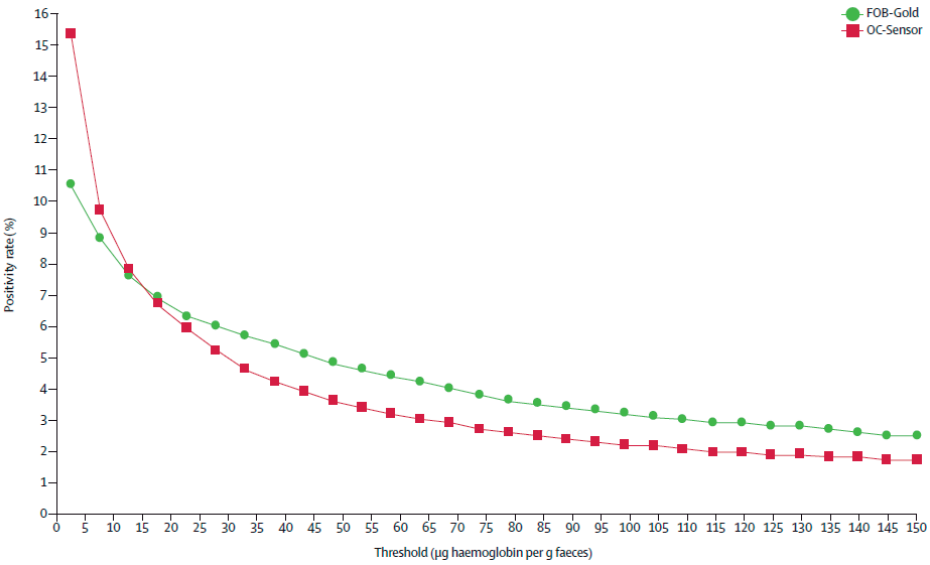
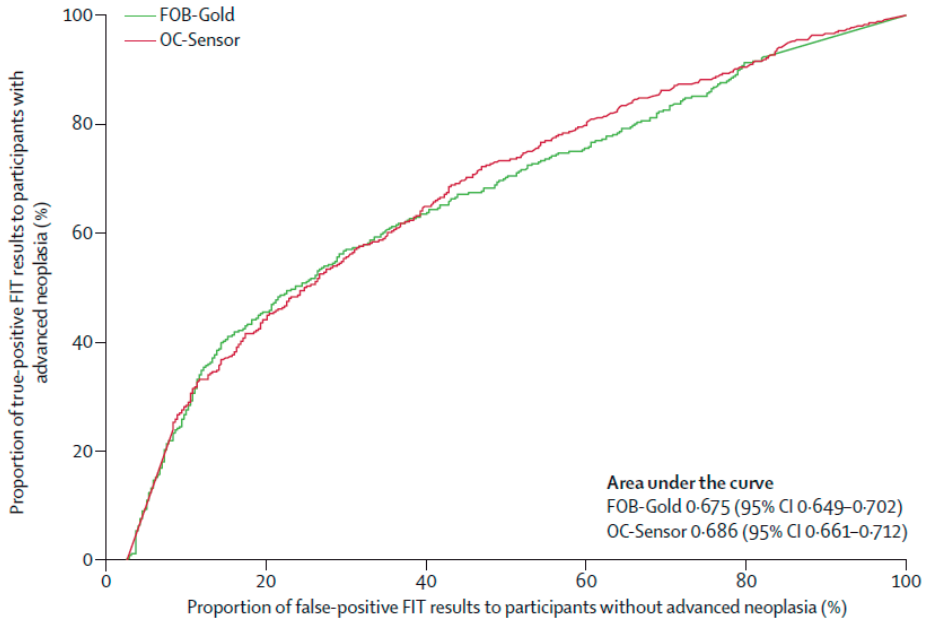


Figure 3 Positivity rates of FOB-Gold and OC-Sensor at potential thresholds



positivity rate of 3.5%, to a maximum difference of 2.4% at a positivity rate of 2.0% (95%CI: -2.5% to 7.3%).

Figure 4 Accuracy of FOB-Gold and OC-Sensor in detecting advanced neoplasia at colonoscopy

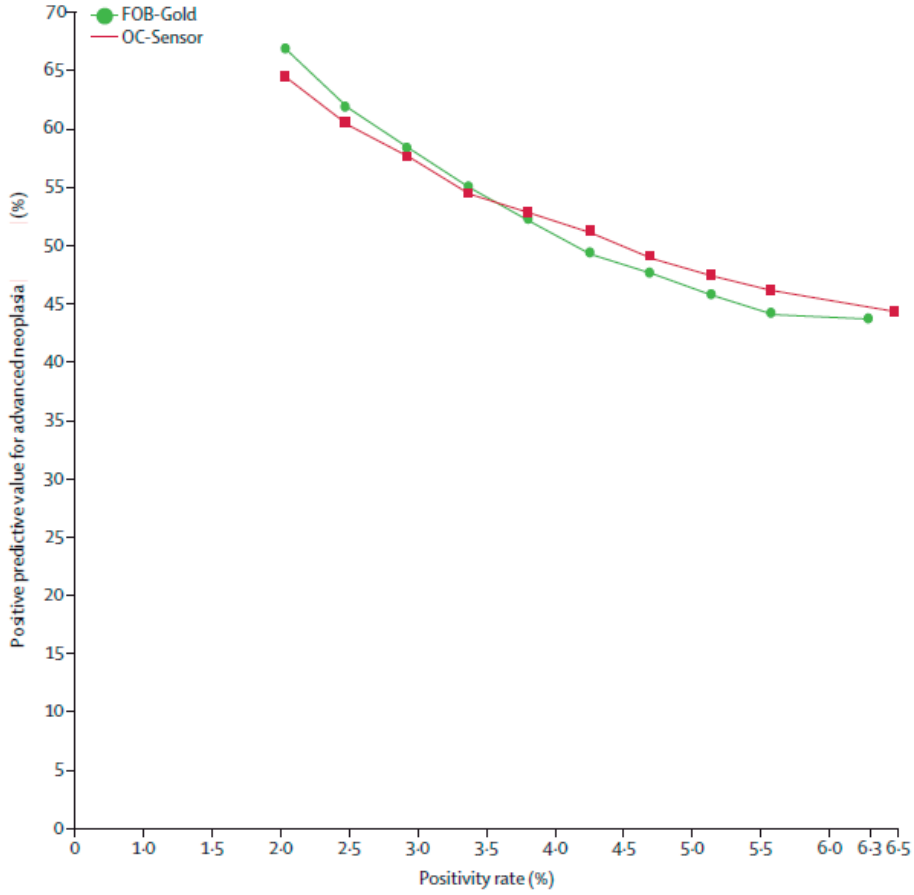


DISCUSSION

We compared reported haemoglobin concentrations with two FITs in over 20,000 participants undergoing colorectal cancer screening who completed both tests in the same bowel movement. Although the haemoglobin concentrations and related positivity rates per haemoglobin threshold within paired samples of FOB-Gold and OC-Sensor tests differed, their capacity to identify participants with advanced neoplasia at colonoscopy was similar when compared at identical positivity rates. This finding implies that FOB-Gold and OC-Sensor are equally accurate in the detection of advanced neoplasia in organised colorectal cancer screening, if the haemoglobin positivity threshold is standardised to yield the same positivity rate.

When interpreting our results, some limitations have to be considered. Because we did not have informed consent for invitees who did not participate in the study, no demographic data were available for these individuals. Participants with two FIT results of less than 15 µg haemoglobin per g faeces were not invited for colonoscopy, so we could not evaluate the false negative rates and the discriminatory performance of FOB-Gold and OC-Sensor in the total study group. We could not calculate estimates of sensitivity and specificity and our estimates of AUC do not reflect the full accuracy of the two tests. About half of the invitees agreed to participate and to complete two tests in this study. This participation is lower than the participation rate of 73% in the national programme, with one test.[8] Although the

Figure 5 Positive predictive value for advanced neoplasia at fixed positivity rates for FOB-Gold and OC-Sensor



study group represents an unselected sample of the intended use population, we cannot exclude selection bias. However, this bias is unlikely to influence our results regarding the comparative performance of the two tests. Although this investigation was a large, paired study, producing valid comparisons, the nature of the Dutch FIT-based colorectal cancer screening programme is such that the positivity rates and predictive values reported here are not unconditionally applicable to other screening settings, populations, and individuals who are not screening naïve. We are aware that manufacturers might adjust their products in the future, which could affect the transferability of our results over time.[16]

This study found a significant difference in reported haemoglobin concentration between two types of FIT, even though participants were requested to sample the two FITs on the same day and in the same bowel movement. The exact reasons for this difference are

not entirely clear. A different method of FIT analysis and calibration, a different buffer composition, antibody specificity, or the settings of the analyser, such as maximum concentration reported (maximal 200 µg haemoglobin per g faeces for OC-Sensor and no maximum for FOB-Gold), might have affected the precise differences in haemoglobin distributions. These differences and the fact that the precision of FIT is limited at low haemoglobin concentrations could also explain the wider range of haemoglobin concentrations in the lower limits of OC-Sensor in contrast to FOB-Gold. The limit of detection for FOB-Gold is reported by the manufacturer to be 15 ng haemoglobin per mL buffer (approximately 2.5 µg haemoglobin per g faeces) versus 20 ng haemoglobin per mL buffer (approximately 4 µg haemoglobin per g faeces) for OC-Sensor.[2] As recently discussed by Fraser and collegues, statements about the accuracy of FITs at low faecal haemoglobin concentrations should therefore be interpreted with caution, as well as statements based on very detailed FIT haemoglobin concentration results.[17]

The difference in haemoglobin concentrations between FITs could also be attributed to the different sampling devices and techniques of the FIT. As instructed by the manufacturer, the FOB-Gold probe should be inserted into the faeces at four different places, whereas the OC-sensor should be scraped through the faeces. Blood, if present, is usually not evenly distributed through the faeces, which might influence the amount of haemoglobin measured by each FIT. Therefore, differences in haemoglobin distribution and positivity rate are also probable if two of the same FIT assays are used to sample the same faces sample.

The selected positivity rate varies across countries and even across screening regions, typically aiming at a high diagnostic yield while limiting the proportion of negative colonoscopies, and taking resource constraints, capacity, and costs into account.[18] Several studies have evaluated FIT positivity thresholds and reported on the corresponding positivity rate for in different screening programmes. In the USA, 20 µg haemoglobin per g faeces is mostly used; [19] in France 20-30 µg haemoglobin per g faeces was selected; [20] in Germany the selected threshold ranges between 9 and 25 µg haemoglobin per g faeces; [21] and in Thailand, 25 µg haemoglobin per g faeces was recommended because colonoscopy capacity is limited.[22] In Scotland, because of scarce colonoscopy capacity, a threshold of 80 µg haemoglobin per g faeces was chosen, resulting in a low FIT positivity rate of 2.4%, similar to the positivity rate with the previously used guaiac faecal occult blood test.[3] In England, a threshold of 120 µg haemoglobin per g faeces is considered for implementation. whereas Wales is expected to launch its screening programme with a threshold of 150 µg haemoglobin per g faeces.[23, 24]

The accuracy of these two widely used FITs has previously been shown to be similar at a positivity threshold of 15 µg haemoglobin per g faeces.[6] Our study supports this equal accuracy for other thresholds, provided the positivity rate is adjusted for. The positivity rate was similar at 15 µg haemoglobin per g faeces but would have differed if a higher or a lower positivity threshold had been selected for the two FITs. These results suggest a stepwise process for selecting the FIT positivity threshold in a screening programme. The developers of such a programme could first select the desired positivity rate or predictive value. Then, data from this study or from similar studies that reflect the intended setting and population could be used to select a preliminary positivity threshold. In a pilot study, the developers could invite screening participants for colonoscopy on the basis of a slightly lower threshold. Then the positivity rate and positive predictive value could be evaluated for the preliminary threshold. The threshold could then be modified, if necessary, on the basis of the desired positivity rate or predictive value. The opposite approach, starting with a higher threshold and increasing the sensitivity from there, could be an alternative approach. Participation rates in the pilot study could be used to further evaluate the burden on colonoscopy services, potentially leading to further modifications of the threshold. In the Netherlands, for example, the threshold was raised from 15 µg to 47 µg haemoglobin per g faeces 6 months after the implementation of the national programme, because thepositivity rate was higher than expected, as the oldest screenees were invited first.[8]

At present, more stratified or personalised screening strategies are being explored with positivity thresholds that differ across population subgroups.[25-28] A higher incidence of advanced neoplasia in men than in women and in individuals eligible for screening of an older age than of a younger age, for example, leads to higher positivity rates but also to higher proportion of missed diagnosis in these subgroups, if the same threshold is used for all individuals. Changing the threshold for specific subgroups could than lead to positive predictive values that are similar across subgroups, if so desired. However, before implementation of sex or age specific thresholds is considered, the actual consequences should first be evaluated in the intended screening setting and population.

This large trial, implemented within an organised colorectal cancer screening programme, showed that despite the significant differences in the distributions of reported haemoglobin, FOB-Gold and OC-Sensor can be considered exchangeable based on their performance in detecting advanced neoplasia if they are standardized at the same positivity threshold.

Acknowledgements

The authors would like to thank the Netherlands Organization for Health Research and Development of the Dutch Ministry of Health (ZonMw) for funding, and all participants for their efforts to perform two tests and a questionnaire. The authors thank all involved coworkers of the Foundation of Population Screening South-West (Bevolkingsonderzoek Zuid-West), and all involved co-workers at Daklapack packaging for their contributions to the study. Finally, the authors thank Star lab (Rotterdam, The Netherlands), for their efforts in analyzing twice as many FIT compared to the nationwide Dutch screening programme during the study period.

REFERENCES

1. Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). Off J Eur Union 2003:34-38.
2. Halloran SP, Launoy G, Zappa M, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy 2012;44 Suppl 3:SE65-87.
3. Steele RJ, McDonald PJ, Digby J, et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity. United European Gastroenterol J 2013;1:198-205.
4. Dancourt V, Lejeune C, Lepage C, et al. Immunochemical faecal occult blood tests are superior to guaiac-based tests for the detection of colorectal neoplasms. Eur J Cancer 2008;44:2254-8.
5. Passamonti B, Malaspina M, Fraser CG, et al. A comparative effectiveness trial of two faecal immunochemical tests for haemoglobin (FIT). Assessment of test performance and adherence in a single round of a population-based screening programme for colorectal cancer. Gut 2018;67:485-496.
6. Wieten E, de Klerk CM, van der Steen A, et al. Equal Accuracy of 2 Quantitative Fecal Immunochemical Tests in Detecting Advanced Neoplasia in an Organized Colorectal Cancer Screening Program. Gastroenterology 2018.
7. Grobbee EJ, van der Vlugt M, van Vuuren AJ, et al. A randomised comparison of two faecal immunochemical tests in population-based colorectal cancer screening. Gut 2016;66:1975-1982.
8. Toes-Zoutendijk E, van Leerdam ME, Dekker E, 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;152:767-775.
9. Fraser CG, Allison JE, Halloran SP, et al. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. J Natl Cancer Inst 2012;104:810-4.
10. National Institute for Public Health and the Environment (RIVM). Exclusion criteria colonoscopy (Dutch) http://www.rivm.nl/Documenten_en_publicaties/Professioneel_Praktisch/Richtlijnen/Preventie_Ziekte_Zorg/Darmkanker/Exclusiecriteria_coloscopie (date accessed 20th June 2017), 2013.
11. Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 2007;29:19-24.
12. 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.
13. Good PI. A Practical Guide to Resampling Methods for Testing Hypotheses. New York: Springer Science & Business Media, 2013.
14. Fagerland MW, Lydersen S, Laake P. Recommended tests and confidence intervals for paired binomial proportions. Stat Med 2014;33:2850-75.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
16. Fraser CG. Comparison of quantitative faecal immunochemical tests for haemoglobin (FIT) for asymptomatic population screening. Translational Cancer Research 2016:S916-S919.
17. Fraser CG, Benton SC. Detection capability of quantitative faecal immunochemical tests for haemoglobin (FIT) and reporting of low faecal haemoglobin concentrations. Clin Chem Lab Med 2018.

18. Symonds EL, Osborne JM, Cole SR, et al. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. J Med Screen 2015;22:187-93.
19. Allison JE, Fraser CG, Halloran SP, et al. Population screening for colorectal cancer means getting FIT: the past, present, and future of colorectal cancer screening using the fecal immunochemical test for hemoglobin (FIT). Gut Liver 2014;8:117-30.
20. Hamza S, Dancourt V, Lejeune C, et al. Diagnostic yield of a one sample immunochemical test at different cut-off values in an organised screening programme for colorectal cancer. Eur J Cancer 2013;49:2727-33.
21. Brenner H, Werner S. Selecting a Cut-off for Colorectal Cancer Screening With a Fecal Immunochemical Test. Clin Transl Gastroenterol 2017;8:e111.
22. Aniwan S, Ratanachu Ek T, Pongprasobchai S, et al. The Optimal Cut-Off Level of The Fecal Immunochemical Test For Colorectal Cancer Screening in a Country with Limited Colonoscopy Resources: A Multi-Center Study from Thailand. Asian Pac J Cancer Prev 2017;18:405-412.
23. Murphy J, Halloran S, Gray A. Cost-effectiveness of the faecal immunochemical test at a range of positivity thresholds compared with the guaiac faecal occult blood test in the NHS Bowel Cancer Screening Programme in England. BMJ Open 2017;7:e017186.
24. Cooper JA, Parsons N, Stinton C, et al. Risk-adjusted colorectal cancer screening using the FIT and routine screening data: development of a risk prediction model. Br J Cancer 2018;118:285-293.
25. Grobbee EJ, Wieten E, Hansen BE, et al. Fecal immunochemical test-based colorectal cancer screening: The gender dilemma. United European Gastroenterol J 2017;5:448-454.
26. Chen SL, Hsu CY, Yen AM, et al. Demand for Colonoscopy in Colorectal Cancer Screening Using a Quantitative Fecal Immunochemical Test and Age/Sex-Specific Thresholds for Test Positivity. Cancer Epidemiol Biomarkers Prev 2018;27:704-709.
27. McDonald PJ, Strachan JA, Digby J, et al. Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. Clin Chem Lab Med 2011;50:935-40.
28. Alvarez-Urturi C, Andreu M, Hernandez C, et al. Impact of age- and gender-specific cut-off values for the fecal immunochemical test for hemoglobin in colorectal cancer screening. Dig Liver Dis 2016.

Supplementary Table 1 Differences in positivity rates of FOB-Gold and OC-Sensor at potential thresholds

Threshold (µg Hb/g)	FOB-Gold positive (n)	FOB-Gold positive (%)	OC-Sensor positive (n)	OC-Sensor positive (%)	Absolute difference (%)	95%CI
5	2210	10.5	3219	15.3	-4.8	-5.2 to -4.4
10	1864	8.8	2051	9.7	-0.9	-1.2 to -0.6
15	1597	7.6	1651	7.8	-0.3	-0.5 to 0.0
20	1452	6.9	1418	6.7	0.2	-0.1 to 0.4
25	1336	6.3	1250	5.9	0.4	0.2 to 0.7
30	1261	6.0	1096	5.2	0.8	0.6 to 1.0
35	1194	5.7	974	4.6	1.0	0.8 to 1.3
40	1135	5.4	893	4.2	1.2	0.9 to 1.4
45	1075	5.1	821	3.9	1.2	1.0 to 1.4
50	1017	4.8	763	3.6	1.2	1.0 to 1.4
55	975	4.6	721	3.4	1.2	1.0 to 1.4
60	921	4.4	679	3.2	1.2	0.9 to 1.4
65	877	4.2	640	3.0	1.1	0.9 to 1.3
70	834	4.0	607	2.9	1.1	0.9 to1.3
75	798	3.8	578	2.7	1.9	1.7 to 2.1
80	758	3.6	553	2.6	1.0	0.8 to 1.2
85	738	3.5	523	2.5	1.0	0.8 to 1.2
90	716	3.4	511	2.4	1.0	0.8 to 1.2
95	690	3.3	490	2.3	1.0	0.8 to 1.1
100	668	3.2	472	2.2	0.9	0.8 to 1.1
105	650	3.1	454	2.2	0.9	0.8 to 1.1
110	633	3.0	439	2.1	0.9	0.8 to 1.1
115	619	2.9	426	2.0	0.9	0.7 to 1.1
120	603	2.9	412	2.0	0.9	0.7 to 1.1
125	589	2.8	403	1.9	0.9	0.7 to 1.1
130	580	2.8	392	1.9	0.9	0.7 to 1.1
135	566	2.7	387	1.8	0.9	0.7 to 1.0
140	551	2.6	376	1.8	0.8	0.7 to 1.0
145	535	2.5	365	1.7	0.8	0.7 to 1.0
150	518	2.5	357	1.7	0.8	0.6 to 0.9

Supplementary Table 2 Differences in positive predictive value for advanced neoplasia at fixed positivity rates for FOB-Gold and OC-Sensor

Positivity rate (%)	**FIT positive (n)**	**AN FOB-Gold (n)**	**(%)**	**AN OC-Sensor (n)**	**(%)**	**Absolute difference (%)**	95% CI
2.0	422	283	67.1	273	64.7	2.4	-2.5 to 7.3
2.5	527	327	62.1	320	60.7	1.3	-2.8 to 5.4
3.0	632	370	58.5	365	57.8	0.8	-2.8 to 4.3
3.5	738	407	55.1	403	54.6	0.5	-2.6 to 3.7
4.0	843	440	52.2	446	52.9	-0.7	-3.4 to 2.0
4.5	949	469	49.4	487	51.3	-1.9	-4.4 to 0.6
5.0	1056	503	47.6	517	49.0	-1.3	-3.7 to 1.0
5.5	1156	529	45.8	548	47.4	-1.6	-3.9 to 0.6
6.0	1265	559	44.2	583	46.1	-1.9	-4.0 to 0.2
6.3	1331	583	43.8	-	-	-	-
6.5	1363	-	-	604	44.5	-	-

6

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

Clinical Gastroenterology and Hepatology (in press)

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

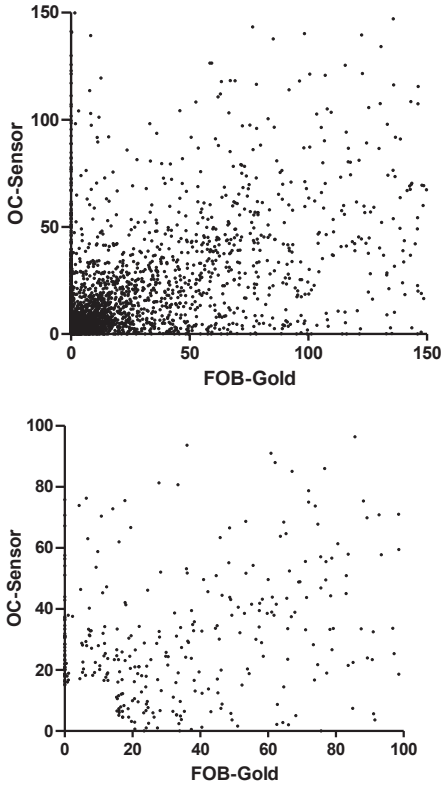
Of the 42,179 screening invitees, 22,111 (52%) returned at least one test and 21,078 (50%) returned both FOB-Gold and OC-sensor. Of the participants who returned both FIT assays, 10,589 (50%) were male. In total, 2,046 (10%) were referred for colonoscopy (one or two positive FIT results).

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%).

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

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

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.

ACKNOWLEDGEMENTS

The authors would like to thank the Netherlands Organization for Health Research and Development of the Dutch Ministry of Health (ZonMw) for funding, and all participants for their efforts to perform two tests and a questionnaire. The authors thank all involved co-workers of the Foundation of Population Screening South-West (Bevolkingsonderzoek Zuid-West), and all involved co-workers at Daklapack packaging for their contributions to the study. Finally, the authors thank Star lab (Rotterdam, The Netherlands), for their efforts in analyzing twice as many FITs compared to the nationwide Dutch screening program during the study period.

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

7

Effects of increasing screening age and fecal hemoglobin cut-off concentration in a colorectal cancer screening program

E. Wieten*, E.H. Schreuders*, S.A.V. Nieuwenburg, B.E. Hansen, I. Lansdorp-Vogelaar, E.J. Kuipers, M.J. Bruno, M.C.W. Spaander.

*shared first author

Clinical Gastroenterology and Hepatology 2016

ABSTRACT

Background & Aims

Several countries have implemented programs to screen for colorectal cancer (CRC) using the fecal immunochemical test (FIT). These programs vary considerably in age of the population screened and the cut-off concentration of fecal hemoglobin (Hb) used to identify candidates for further evaluation; these variations are usually based a country's colonoscopy resources. We calculated how increasing the Hb cut-off concentration and screening age affects colonoscopy yield, missed lesions, and demand.

Methods

We collected data from 10,008 average-risk individuals in The Netherlands, 50–74 years old, who were invited for a FIT in the first round of a population-based CRC screening program from November 2006 through December 2008. Fecal samples were collected and levels of Hb were measured using the OC-sensor Micro analyzer; concentrations ≥10 μg Hb/g feces were considered positive. Subjects with a positive FIT were scheduled for colonoscopy within 4 weeks. Logistic regression analysis was performed to evaluate the association between age and detection of advanced neoplasia.

Results

In total, 5986 individuals (62%) participated in the study; 503 had a positive test result (8.4%). Attendance, positive test results, detection advanced neoplasia, and the FIT's positive predictive value (PPV) all increased significantly with age ($P<.001$). Detection of advanced neoplasia ranged from 1.3% in the youngest age group to 6.2% in the oldest group; the PPV value of the FIT was 26% in the youngest group and 47% in the oldest group. Increasing the starting age of invitees from 50–74 years to 55–74 years reduced the proportion of subjects who underwent colonoscopy evaluation by 14% and resulted in 9% more subjects with advanced neoplasia being missed. Increasing the cut-off concentration from 10 to 15 μg Hb/g feces reduced the proportion of subjects who underwent colonoscopy evaluation by 11% and resulted in 6% of advanced neoplasia being missed.

Conclusion

In an analysis of an average-risk screening population in The Netherlands, we found that detection of advanced neoplasia by FIT increases significantly with age and fecal Hb cut-off concentration. Increasing the cut-off concentration or screening age reduces the numbers of patients who undergo colonoscopy evaluation in FIT-based CRC screening programs. Our findings provide insight in these effects per age category and cut-off concentration, and the consequences, in terms of missed lesions.

INTRODUCTION

Colorectal cancer (CRC) is the second most common cause of cancer-related mortality in the Western world.[1] Screening with guaiac-based fecal occult blood tests (gFOBT) can reduce CRC-related mortality.[2] The gFOBT is now gradually replaced by fecal immunochemical test (FIT) for hemoglobin because of its superior adherence and accuracy.[3] Quantitative FIT furthermore offers the opportunity of selecting a specific cut-off fecal Hb concentration used to identify candidates for further evaluation; that provides an optimal match between screening population and available financial and endoscopy resources.

In recent years, several countries have implemented a FIT-based nationwide CRC screening program.[4] Cut-off concentration and age of the population screened vary between countries, often tailored to available financial resources and colonoscopy capacity.[5, 6] For example, organized FIT screening is offered to 55-75 year-olds in the Netherlands, using a positivity cut-off of 47 µg Hb/g feces, whereas in the United Kingdom 60-74 year-olds are invited for FIT screening and a cut-off concentration of 20 µg Hb/g feces is used. A high cut-off and narrow screening age range result in a low positivity rate and consequently low colonoscopy demand.[7] However, this comes at the cost of a decrease in detection rate of advanced neoplasia (AN).[7]

Previous studies showed that the prevalence of AN, defined as colorectal cancer and advanced adenomas, increases with age in individuals undergoing a screening colonoscopy.[8, 9] Also, fecal Hb concentrations determined by FIT tend to increase with age, and higher positivity rates and detection rates are found in elder screenees compared to younger screenees.[10-12] Therefore, age partitioned cut-offs for fecal Hb concentration may be warranted if AN detection rate increases relatively slower than positivity rate.

The aim of this study was to assess positivity rates and detection rates of FIT in different age categories and to assess how this relates to the positive predictive value (PPV) in a population-based CRC screening program. Our secondary aim was to estimate the effect of increasing the cut-off concentration and screening age on the numbers of patients who undergo colonoscopy and AN detection and miss rate.

METHODS

Study population

This study compromises the first round of a population-based organized CRC screening program (CORERO-I) by means of FIT, of which the methods and primary results have been described elsewhere.[13-15] In short, 10,008 CRC screening-naïve individuals aged 50-74 years living in region Rotterdam-Rijnmond in the South-West of the Netherlands were randomly selected and invited. We excluded individuals who met one of the exclusion criteria (a

history of inflammatory bowel disease or CRC; colonoscopy, sigmoidoscopy, or barium contrast enema within the previous three years; inability to give informed consent) or who died or moved away. Recruitment took place between November 2006 and December 2008.

For the purpose of this study, we assessed rates of attendance, test positivity, detection of AN, and PPV in the following five age categories; 50-54, 55-59, 60-64, 65-69 and 70-74 years. In these age groups, we also calculated differences in the diagnostic yield of FIT, number needed to screen and number needed to scope to detect one case with AN. We further assessed if the PPV differed with age, when corrected for the confounders gender, socioeconomic status and fecal Hb concentration. We calculated the effect of offering CRC screening to later ages with steps of 5 years on the numbers of patients who undergo colonoscopy and number of detected and missed AN. We also assessed the effect of increasing cut-off concentrations on these parameters. Finally, the effect of these screening strategies was converted into a risk ratio with 95% CI, i.e. the percentage reduction of those who undergo colonoscopy was divided by the percentage of missed advanced neoplasia. This ratio explains the relative decrease in colonoscopy demand per percentage lesion missed advanced neoplasia.

Intervention and follow-up evaluation

One FIT (OC-sensor, Eiken Chemical, Tokyo, Japan) was sent by mail to collect a single sample of one bowel movement. Participants returned the FIT and an informed consent form at ambient temperature by freepost to the Gastroenterology & Hepatology laboratory of the Erasmus Medical Centre, Rotterdam, the Netherlands. The test was analyzed on the OC-sensor Micro system (Eiken, Japan) and considered positive at a fecal Hb concentration of ≥10 μg Hb/g feces (≥ 50 ng Hb/mL) (Appendix 1).

Subjects with a positive FIT were scheduled for colonoscopy within four weeks and subjects with a negative FIT were referred back to the screening program. All colonoscopies were done by experienced gastroenterologists. Removed polyps were evaluated by expert gastrointestinal pathologists. Patients with a positive colonoscopy entered a surveillance program, whereas subjects with a negative colonoscopy were considered not to require FIT screening for 10 years.[16]

Definitions

Attendance rate was calculated by dividing the number of eligible participants by all eligible subjects (all invitees minus the excluded clients). Positivity rate was defined as the proportion of positive tests in participants with an analyzable test-result. Detection rate was defined as those with AN or CRC relative to all participants with an analyzable test. Advanced neoplasia included CRC and advanced adenomas. 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 person, the screenee was classified according to the most advanced lesion. The PPV compromised all screenees diagnosed with AN or CRC proportionally to screenees with a positive FIT who underwent a colonoscopy. The diagnostic yield of FIT per 10,000 eligible invitees was defined as screenees with AN relative to all eligible invitees. Number needed to scope was calculated as the number of colonoscopies needed to find one screenee with AN. Number needed to screen describes the number of complete FITs needed to find one case with AN.

Statistical analyses

Comparisons of continuous variables were performed using the Mann–Whitney U-test. Categorical variables with two categories were compared using the $\chi 2$ test and those with multiple categories with binary logistic regression analyses. Equality of fecal Hb concentration distributions between age categories was tested with the Kruskal-Wallis test. Attendance rate, positivity rate, detection rate and PPV were described as proportions with 95% confidence intervals (CI).

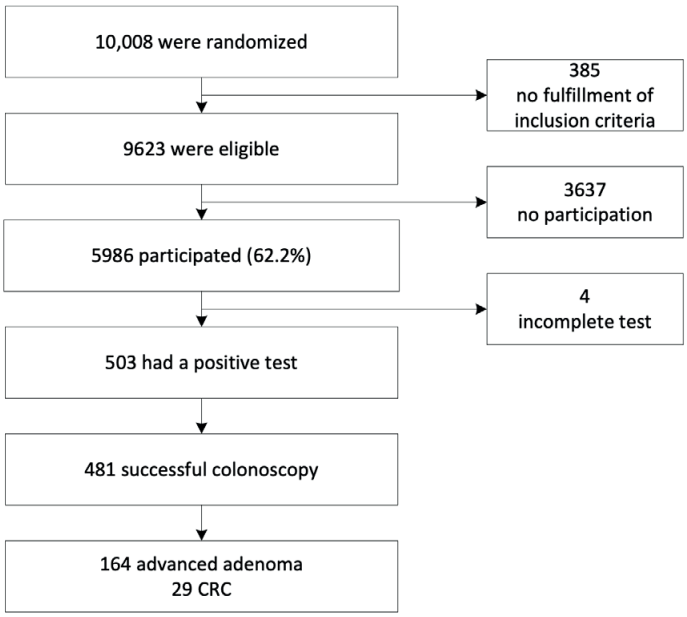
To estimate significant differences in PPV for all ages, a multivariate binary regression analysis was performed, with age included as continuous variable. First, univariate binary logistic regression analyses were performed to determine the independent association of multiple variables (sex, age, social economic status, fecal Hb concentration) with the PPV of AN. In these analyses the PPV was used as outcome variable by selecting all participants with an analyzable FIT. In addition, detection of AN was selected as the dependent variable. Subsequently, all univariate significant variables and variables chosen by the clinician's rationale (i.e. gender) were included in a multivariate logistic regression analysis. Interactions were tested between all variables that were included in the multivariate model. Interactions were included in the final model when significant ($P<0.01$). Hosmer and Lemeshow chi-square statistics were used as goodness of fit statistic. The outcome of the final multivariate logistic regression model resulted in a predicted probability of having AN per screenee who had a positive FIT and subsequent colonoscopy. These predicted probabilities of having AN per screenee were depicted in a figure with age as a continuous variable.

The analyses were performed using SPSS V.21 statistical package (SPSS Inc, Chicago, Illinois, USA). All p-values were two-sided and considered significant if $P<0.05$, except for interactions which were considered significant if $P<0.01$.

Ethical approval

All participants signed informed consent. The study was approved by the Dutch Ministry of Health (2006/02WBO). The invitation letters and information brochures were approved by the Institutional Review Board of the Erasmus MC (MEC-2005-264). All authors had access to the study data and reviewed and approved the final manuscript.

Figure 1 Trial profile



CRC: Colorectal cancer

RESULTS

Participant characteristics

The trial profile is summarized in Figure 1. Of the 10,008 individuals invited, 385 subjects were excluded from analyses, due to various reasons as described previously.[13-15] In total, 5,986 (62%) attended screening, of which 5,982 had an analyzable screening test. A total of 503 (8.4%) screenees had a positive test at a cut-off concentration of ≥10 µg Hb/g feces of which 481 (96%) underwent colonoscopy.

Test characteristics per age

Test results per age category are given in Table 1. Screening participants had a median age of 61 years and 48% of the participants was male. Attendance rate and positivity rate increased significantly with age (P<.001). Detection rates of AN and CRC increased significantly with age as continuous variable (P<.001), and ranged from 1.3% for AN in the youngest age category of 50 to 54 years old, to 6.2% in the eldest age category of 70 to 74 years. The PPV for AN per age category ranged from 26% to 47%. The number needed to screen and number needed to scope to detect one case with AN ranged from 138 to 26 and 3.9 to 2.1, respectively. The diagnostic yield of AN per 10,000 eligible invitees ranged from 73 in the 50 to 54 year-old to 392 in the 70 to 74 year-old.

Table 1 Test characteristics of FIT for five different age categories in a colorectal screening program.

Age categories	Eligible invitees	Attendance rate	Positivity rate	Underwent colonoscopy	Detection rate		PPV		Number needed to screen to detect one person with AN	Number needed to scope detect one person with AN	Diagnostic yield of AN per 10,000
					AN	CRC	AN	CRC			
years	n	n (%)	n (%)	n (%)	n (%)	n (%)	%	%			
50-54	2,343	1,343 (57)	68 (5.1)	66 (97)	17 (1.3)	1 (0.07)	25.8	1.5	138	3.9	73
55-59	2,381	1,467 (62)	106 (7.2)	99 (93)	41 (2.8)	8 (0.55)	41.4	8.1	58	2.4	172
60-64	2,149	1,419 (66)	119 (8.4)	113 (95)	41 (2.9)	5 (0.35)	36.3	4.4	52	2.8	190
65-69	1,577	1,015 (64)	110 (10.8)	106 (96)	48 (4.7)	6 (0.59)	45.3	5.7	33	2.2	304
70-74	1,173	742 (63)	100 (13.5)	97 (97)	46 (6.2)	9 (1.21)	47.4	9.3	26	2.1	392
Total	9,623	5,986 (60)	503 (8.4)	481 (96)	193 (3.2)	29 (0.50)	40.1	6.0	50	2	201

FIT: fecal immunochemical test; PPV: positive predictive value; AN: advanced neoplasia; CRC: colorectal cancer.

Test characteristics per cut-off

Fecal Hb concentration ranged from 0 to 921 µg Hb/g feces in participants with an analyzable screening test. Fecal Hb concentration increased significantly with age as a continuous variable (p<.001) and was significantly higher in men than in women (p<.001). Positivity rate decreased from 8,4% to 6,2%, 4,4% and 3,8% when the cut-off was increased from ≥10 µg Hb/g feces to sequentially ≥20 µg Hb/g feces, ≥30 µg Hb/g feces, and ≥40 µg Hb/g feces.

Regression model

Univariate logistic regression analyses showed that an increase in age (*P*=.003) and increase in fecal Hb concentration (*P*<.001) were both associated with a higher PPV (Table 2). Age remained significantly related to PPV in the multivariate logistic regression analysis, with an odds ratio (OR) of 1.53 per 10 years (95% CI 1.13-2.07), when corrected for fecal Hb concentration and gender. There were no significant interactions in this model and no significant differences in frequencies between the observed values and the predicted values (Goodness-of-Fit; *P*=.276). The predicted probability of having AN for participants with a positive FIT per age, corrected for gender and fecal Hb concentration, was depicted in Figure 2 for different fecal Hb concentration subgroups.

Table 2 Univariate and multivariate logistic regression analyses of factors associated with the detection of advanced neoplasia in a FIT based CRC screening program.

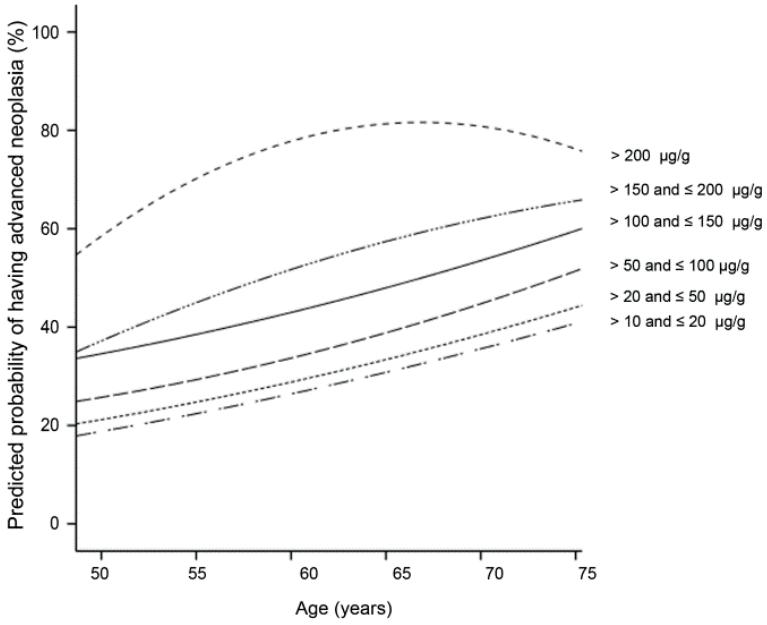
	Univariate		**Multivariate**	
	Odds ratio (95% CI)	**p-value**	**Odds ratio** (95% CI)	**p-value**
Sex male	1.24 (0.85-1.81)	.27	1.12 (0.74-1.68)	.60
Age (per ten years increase)	1.53 (1.15-2.03)	.003	1.53 (1.13-2.07)	.005
Socio-economic status (SES)		.43		
Low	*Reference*			
Middle	0.90 (0.54-1.48)			
High	0.76 (0.51-1.15)			
Fecal Hb concentration (per 10 µg Hb/g feces increase)	1.07 (1.05-1.09)	<.001	1.07 (1.05-1.09)	<.001

FIT: fecal immunochemical test; CRC: colorectal cancer.

Increasing the screening starting age and cut-off

Colonoscopy demand and detection rate of AN per age category and cut-off concentration are shown in Figure 3a and Figure 3b. Increasing the starting age from 50 to 55 years resulted in a total decrease in colonoscopy demand of 14%, however at the expense of missing 9% of AN (Table 3). If solely the cut-off was increased from 10 to 12,5 µg Hb/g feces, this resulted in a decreased colonoscopy demand of 11%, at the expense of missing

Figure 2 The predicted probability of having advanced neoplasia displayed for screenees with a positive FIT per age and different fecal hemoglobin concentrations* (corrected for gender)



*microgram hemoglobin per gram feces
FIT: fecal immunochemical test.

7% AN. Thus in both strategies, for every 1.6% decrease in subjectss who undergo colonoscopies, 1% of screenees with AN were missed, resulting in a 1.6 ratio.

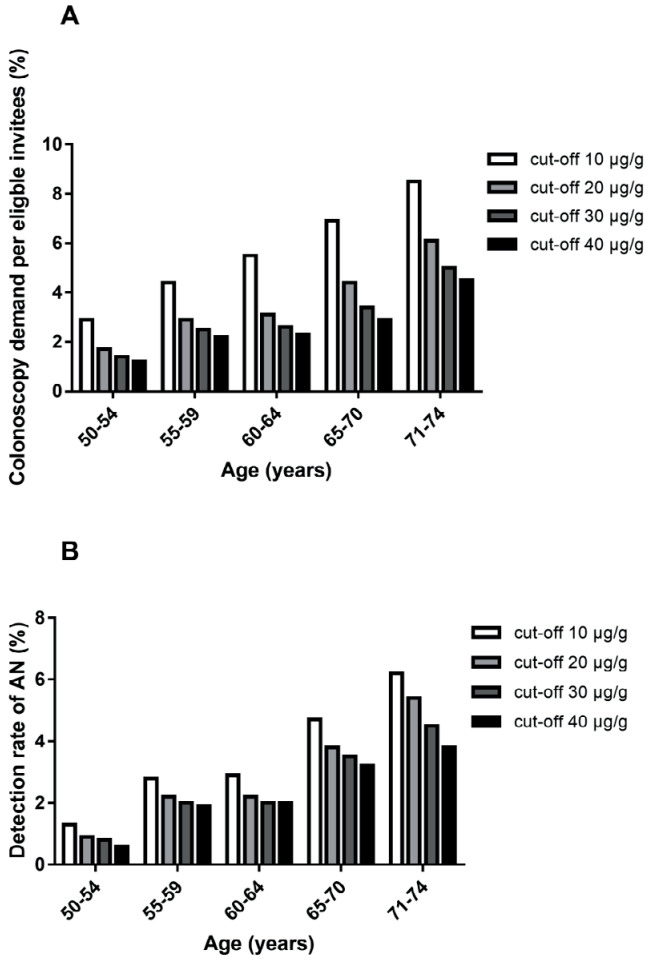
Screening strategies with age-specific cut-off concentrations were also assessed on these outcomes (Table 3). The highest benefit of the reduction of numbers of patients who undergo colonoscopy relative to the loss in AN detection was achieved by strategy 6 and, while strategy one and five resulted in the lowest decrease of AN detection rate.

Total colonoscopies needed and detected lesions are shown, compared to the reference screening strategy (first round, screen age 50-74 years, cutoff 10 µg Hb/g feces). Screening strategies are altered by 1. increasing the starting age and 2. increasing cut-offs and 3. Combinations of age and cut-off alterations FIT: fecal immunochemical test; Hb: Hemoglobin; AN: Advanced Neoplasia; CRC: colorectal cancer; µg: microgram; g: gram; PPV: positive predictive value.

DISCUSSION

In this study we showed that there were substantial differences in diagnostic yield of FIT between age groups. In this population-based CRC screening cohort, FIT positivity

Figure 3a) Colonoscopy demand and **b)** detection rate of advanced neoplasia (AN) per age category and cut-off concentration in µg Hb/gram feces.



rates, detection rates and the PPV all significantly increased with age. Both increasing the screening starting age and increasing the cut-off concentration resulted in a substantial reduction in colonoscopy demand.

Currently, cut-off concentrations and age of the population screened varies between countries with a FIT-based screening program.[4] FIT-based screening with more tailored approaches based on sex, age or risk factors have been suggested in several studies.[11, 12] We calculated the effect of increasing the screening starting age and cut-off concentration on colonoscopy demand and number of detected and missed AN. We showed that both

Table 3 Effects of FIT-based CRC screening strategies on the numbers of patients who undergo colonoscopy and missed lesions in a screening population of 10,000 eligible invitees.

Number strategy	Screening strategies	Total colonoscopies needed per 10,000 eligible invitees	Detected lesions per 10,000 eligible invitees		Ratio of percentage decrease in colonoscopies needed per percentage AN missed
	Age range, cut-off (µg Hb/g feces)	N (%*)	AN (%*)	CRC (%*)	(95% CI)
	50-74, 10 (reference)	*500*	*201*	*30*	
1	55-74, 10	431 (-14)	183 (-9)	29 (-3)	1.6 (0.94 – 2.52)
2	60-74, 10	328 (-34)	141 (-30)	21 (-30)	1.2 (0.90 – 1.47)
3	65-74, 10	211 (-58)	98 (-51)	15 (-50)	1.1 (0.97 – 1.32)
4	70-74, 10	99 (-80)	48 (-76)	9 (-70)	1.1 (0.96 – 1.15)
5	50-74, 12.5	447 (-11)	188 (-7)	28 (-7)	1.6 (0.91 – 2.93)
6	50-74, 15	374 (-25)	175 (-13)	28 (-7)	1.9 (1.32 – 2.88)
7	50-74, 20	316 (-37)	160 (-20)	28 (-7)	1.9 (1.34 – 2.43)
8	50-74, 50	202 (-60)	119 (-41)	21 (-30)	1.5 (1.21 – 1.75)
9	50-74, 100	134 (-73)	93 (-54)	16 (-47)	1.4 (1.19 – 1.57)
10	50-54, 30; 55-59, 25; 60-64, 20; 65-69, 15; 70-74, 10	347 (-31)	166 (-18)	29 (-3)	1.8 (1.23 – 2.44)
11	50-54, 10; 55-59, 15; 60-64, 20; 65-69, 25; 70-74, 30	334 (-33)	160 (-21)	27 (-10)	1.6 (1.21 – 2.20)

*percentage decrease

actions result in a substantial reduction in colonoscopy demand. However, also AN and CRC are missed subsequently. Increasing the screening starting age and increasing the cut-off concentration had different effects on the absolute numbers of screenees who undergo colonoscopy and missed lesions. An interesting finding was however that in relative ratios equal effects were found. For every missed AN in one screenee, both increasing the cut-off concentration and increasing the starting age, resulted in a similar decrease in screenees who undergo colonoscopy.

In our cohort, the PPV for AN increased significantly with age, even when corrected for confounders. This is in line with a recent Spanish study.[17] A likely explanation for the increasing PPV of FIT with age is that a greater proportion of AN occurs in elder persons.[8, 9]

Lower detection rates compared to elder individuals are generally accepted as younger persons have more life-years to gain.[18]

In addition to age, we showed that incremental increase of fecal Hb concentration detected by FIT is associated with an increase in PPV for AN, suggesting (pre)malignant lesions bleed more compared to other lesions. This mechanism is also suggested in previous literature.[17] Differences in detection rates between age groups, sex and fecal Hb concentrations determined by FIT have been described before in FIT-based CRC screening populations.[19-21] In our cohort, sex was not associated with the PPV for AN. This is most likely due to the small absolute numbers of screenees with AN in our study. Literature has shown that FIT detects AN more often in males than females.[11, 22]

The relation between fecal hemoglobin concentrations and age on the predicted probability of having AN was linearly shaped for the lower concentrations and parabolically shaped for the higher concentrations. This parabolic shape is possibly a result of the low numbers of subjects with high fecal Hb concentrations for the youngest and oldest ages. In addition, a prozone effect might have occurred at very high fecal Hb concentrations (>200 µg Hb/g feces). A prozone effect can appear if the fecal Hb concentration exceeds the limit of antigen agglutination.[23] As a result, the actual sample concentration can be higher than the measured Hb values for values >200 µg Hb/g feces. It should be of note that it is hypothesized that FIT is relatively insensitive for the detection of serrated neoplasia.[24] Serrated lesions are thought to bleed less often compared with adenomas.[25-27]

An optimal cut-off concentration or screening age range could not be established in this study. Evidence has been published that in a Western population the optimal cut-off concentration is low (10 µg Hb/g feces) and screening age range is wide (45-80 years).[18] However, such screening programs require the availability of unlimited colonoscopy resources. In our study we showed that alterations in cut-off concentrations and age of the population screened are both good options when facing colonoscopy capacity limitations, dependent on available resources.

The main limitation of our study is the size of the cohort. Since CRC was detected in only few persons, changes in the cut-off concentration or screening starting age had a substantial effect in percentage of missed CRC. Therefore, the effect of increasing the screening starting age and cut-off concentration on the missed CRCs should be carefully interpreted. However, our study size was acceptable to calculate effects of screening alterations on detected and missed AN. Information on age differences in missed lesions in population based FIT screening has been limited until now. This study provides insight into this matter. It should however be taken into account that in our cohort only those

subjects with a cut-off ≥ 10 µg Hb/g feces were offered colonoscopy. Missed lesions below this cut-off could therefore not be evaluated.

In conclusion, increased age is associated with an increase in positivity rates, detection rates and PPV in a FIT-based CRC screening cohort. Our findings give insight in the effect of increasing the screening starting age and cut-off concentration on colonoscopy demand and missed lesions in absolute numbers. When facing colonoscopy capacity problems, these effects can be taken into account. Further research to evaluate the impact of age-tailored cut-offs in multiple screening rounds is needed.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.
2. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev. 2007(1):CD001216.
3. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327-37.
4. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJY, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49.
5. van Hees F, Zauber AG, van Veldhuizen H, Heijnen ML, Penning C, de Koning HJ, et al. The value of models in informing resource allocation in colorectal cancer screening: the case of the Netherlands. Gut. 2015;64(12):1985-97.
6. Steele RJ, McDonald PJ, Digby J, Brownlee L, Strachan JA, Libby G, et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity. United European Gastroenterol J. 2013;1(3):198-205.
7. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100(7):1103-10.
8. 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):1585-89.
9. Ferlitsch M, Reinhart K, Pramhas S, Wiener C, Gal O, Bannert C, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA. 2011;306(12):1352-58.
10. van Rossum LG, van Rijn AF, Laheij RJ, van Oijen MG, Fockens P, Jansen JB, et al. Cutoff value determines the performance of a semi-quantitative immunochemical faecal occult blood test in a colorectal cancer screening programme. Br J Cancer. 2009;101(8):1274-81.
11. McDonald PJ, Strachan JA, Digby J, Steele RJ, Fraser CG. Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. Clin Chem Lab Med. 2012;50(5):935-40.
12. Fraser CG, Rubeca T, Rapi S, Chen LS, Chen HH. Faecal haemoglobin concentrations vary with sex and age, but data are not transferable across geography for colorectal cancer screening. Clin Chem Lab Med. 2014;52(8):1211-6.
13. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JC, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59(1):62-8.
14. van Roon AH, Goede SL, van Ballegooijen M, van Vuuren AJ, Looman CW, Biermann K, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut. 2013;62(3):409-15.
15. van Roon AH, Hol L, Wilschut JA, Reijerink JC, van Vuuren AJ, van Ballegooijen M, et al. Advance notification letters increase adherence in colorectal cancer screening: a population-based randomized trial. Prev Med. 2011;52(6):448-51.
16. Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the

American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. CA Cancer J Clin. 2008;58(3):130-60.
17. Auge JM, Pellise M, Escudero JM, Hernandez C, Andreu M, Grau J, et al. Risk Stratification for Advanced Colorectal Neoplasia According to Fecal Hemoglobin Concentration in a Colorectal Cancer Screening Program. Gastroenterology. 2014;147(3):628-36.e1.
18. Wilschut JA, Hol L, Dekker E, Jansen JB, Van Leerdam ME, Lansdorp-Vogelaar I, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology. 2011;141(5):1648-55 e1.
19. Chen CH, Tsai MK, Wen CP. Extending Colorectal Cancer Screening to Persons Aged 40 to 49 Years With Immunochemical Fecal Occult Blood Test: A Prospective Cohort Study of 513,283 Individuals. Journal of clinical gastroenterology. 2016.
20. Symonds EL, Osborne J, Cole SR, Bampton P, Fraser R, Young GP. Gender differences in faecal haemoglobin concentration. J Med Screen. 2016;23(1):54.
21. Alvarez-Urturi C, Andreu M, Hernandez C, Perez-Riquelme F, Carballo F, Ono A, et al. Impact of age- and gender-specific cut-off values for the fecal immunochemical test for hemoglobin in colorectal cancer screening. Dig Liver Dis. 2016;48(5):542-51.
22. Kapidzic A, van der Meulen MP, Hol L, van Roon AH, Looman CW, Lansdorp-Vogelaar I, et al. Gender differences in fecal immunochemical test performance for early detection of colorectal neoplasia. Clin Gastroenterol Hepatol. 2015;13(8):1464-71.
23. Vaananen P, Tenhunen R. Rapid immunochemical detection of fecal occult blood by use of a latex-agglutination test. Clinical chemistry. 1988;34(9):1763-6.
24. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med. 2014;370(14):1287-97.
25. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology. 2010;138(6):2088-100.
26. East JE, Saunders BP, Jass JR. Sporadic and syndromic hyperplastic polyps and serrated adenomas of the colon: classification, molecular genetics, natural history, and clinical management. Gastroenterol Clin North Am. 2008;37(1):25-46, v.
27. Waldock A, Ellis IO, Armitage NC, Turner DR, Hardcastle JD. Histopathological assessment of bleeding from polyps of the colon and rectum. J Clin Pathol. 1989;42(4):378-82.

8

Fecal immunochemical test-based colorectal cancer screening: the gender dilemma

E.J. Grobbee, E. Wieten, B.E. Hansen, E.M. Stoop, T.R. de Wijkerslooth, I. Lansdorp-Vogelaar, P.M. Bossuyt, E. Dekker, E.J. Kuipers, M.C.W. Spaander

United European Gastroenteroly Journal 2017

ABSTRACT

Background

Despite differences between men and women in incidence of colorectal cancer (CRC) and its precursors, screening programs consistently use the same strategy for both genders.

Objective

To illustrate the effects of gender tailored screening, including the effects on miss rates of advanced neoplasia (AN).

Methods

Participants (50-75 years) in a colonoscopy screening program were asked to complete a fecal immunochemical test (FIT) before colonoscopy. Positivity rates, sensitivity and specificity for detection of AN at multiple cut-offs were determined. Absolute numbers of detected and missed AN per 1000 screenees were calculated.

Results

In total 1,256 underwent FIT and colonoscopy, 51% male (median age 61 years; IQR 56-66) and 49% female (median age 60 years; IQR 55-65). At all cut-offs men had higher positivity rates than women, ranging from 3.8-10.8% versus 3.2-4.8%. Sensitivity for AN was higher in men than women; 40-25% and 35-22%, respectively. More AN were found and missed in absolute numbers in men at all cut-offs.

Conclusion

More AN were both detected and missed in men compared to women at all cut-offs. Gender tailored cut-offs could either level sensitivity in men and women (i.e. lower cut-off in women) or level the amount of missed lesions (i.e. lower cut-off in men).

INTRODUCTION

Colorectal cancer (CRC) is one of the leading causes of cancer-related death in the Western world.[1, 2] Detection of occult blood in feces by guaiac fecal occult blood testing (gFOBT) has been proven to reduce CRC-related mortality.[3] In recent years, fecal immunochemical testing (FIT) has become the preferred method of detecting fecal occult blood for CRC screening. FIT is more sensitive for the detection of CRC and its precursors.[4, 5] Besides, FIT is easier to handle than gFOBT.[6] Consequently, screening participation rates increase.[4, 7, 8] Also, FIT-analysis can be automated and quantitated. Quantitative FITs enable adjustment of cut-off to vary test-characteristics and match demand with available resources, in particular colonoscopy capacity.[9]

Men and women differ with respect to the prevalence of advanced colorectal neoplasia, with men having substantial higher prevalence of advanced adenomas and CRC than women.[10, 11] Repeated biennial gFOBT screening leads to a higher overall mortality reduction in men than in women.[3] In fact, a recent gFOBT-based study showed that the prevalence of colorectal neoplasms was higher in men with a negative test than in women with a positive test.[12] Moreover, male gender seems to be a stronger predictor of CRC than a positive gFOBT.[13] Results from the national gFOBT screening program in Scotland showed a lower proportion of interval cancers for men compared to women.[14] One study, using FIT, showed that men had higher positivity rates, as well as a higher detection rate.[15] However, this study was limited as only FIT-positive (i.e. a fecal hemoglobin concentration >10 µg Hb/g feces) screenees underwent colonoscopy.

As more screening programs are being implemented worldwide, these gender differences become more apparent. Despite these differences, screening programs consistently use same strategies for both genders with regards to cut-off and screening intervals.[10, 16] Even though the use of different cut-offs in men and women would allow tailored screening strategies for each gender and improve CRC screening efficacy.

Most studies on gender differences in CRC screening used gFOBT or FIT with one cut-off for both genders. Results were often based on assessing equal sensitivity of the test for men and women, thereby not taking into account gender differences regarding detection rate and miss rate of lesions in absolute numbers. Therefore, we aimed to illustrate the effect of gender tailored FIT screening including the detection and miss rates of advanced neoplasia.

METHODS

The protocol of this population-based screening pilot (trialregister.nl; identifier NTR3549) has been described previously in detail.[17, 18] All authors had access to the study data and reviewed and approved the final manuscript.

Study population

Between June 2009 and July 2010, 6,600 asymptomatic individuals aged 50-75 years, living in the Amsterdam and Rotterdam regions were randomly selected from regional municipal administration registrations. They were invited for colonoscopy screening as primary screening modality or invited for computed tomography colonography. For the purpose of this manuscript only data of the population undergoing colonoscopy were used.

Individuals with a history of inflammatory bowel disease or CRC, as well as those who had undergone a full colonic examination in the past 5 years, those with an estimated life expectancy of <5 years, and subjects who were unable to give informed consent were excluded from the study. As there was no CRC screening program at the time of the trial in the Netherlands, the target population was screening-naive when first approached.

Fecal occult blood screening and colonoscopy

Eligible subjects who gave informed consent for colonoscopy screening were asked to complete one sample FIT (OC-sensor, Eiken, Japan) before colonoscopy. Participants were instructed to perform FIT at home, within 48 hours before the colonoscopy, but before starting the bowel preparation. No dietary restrictions were given. All patients underwent subsequent colonoscopy by experienced endoscopists. Research staff attended all colonoscopies and prospectively documented colonoscopy quality indicators and data on CRC and polyp detection.

Histology

Experienced pathologists classified all removed lesions as non-neoplastic, serrated polyp, adenoma (tubular, tubulovillous or villous) or carcinoma. Dysplasia was defined as low-grade or high-grade. Advanced adenomas were defined as an adenoma larger than 10 mm, an adenoma with villous histology (>25%) and/or an adenoma with high-grade dysplasia. Advanced neoplasia (AN) included both AA and CRC.

Statistical analysis

All screening participants who completed a FIT and subsequently underwent colonoscopy were included in the analysis. Baseline characteristics were described using descriptive statistics. The Chi-square test was used for comparing proportions of advanced neoplasia between men and women. The Mann-Whitney U test was used for non-parametric

distributions. The sensitivity, specificity, positive and negative predictive value (PPV / NPV), and detection rate (DR) of advanced neoplasia were calculated for the most commonly used cut-offs; 10 (FIT10), 20 (FIT20), 30 (FIT30) and 40 (FIT40) µg Hb/g feces. These values correspond to 50, 100, 150 and 200 ng Hb/ml buffer. Following, sensitivity and specificity for fecal Hb-concentrations for all cutoffs between 0 and 100 µg Hb/g feces were calculated. Absolute numbers of detected and missed AN per 1000 subjects screened were calculated for men and women.

RESULTS

Baseline characteristics and colonoscopy outcome

In total 1,256 invitees underwent FIT and colonoscopy, 638 men and 618 women. Men (61 years, IQR 56-66) were slightly older than women (60 years, IQR 55-65). Gender-specific findings at colonoscopy are described in Table 1. AN detection rate was slightly higher in men than in women, 10.6% (68/638) versus 8.3% (51/618) (p=0.146). CRC was detected in 5 (0.8%) men and in 3 (0.5%) women. No differences between men and women were seen in location of AN or number of AN per participant. In subjects with CRC, the median fecal hemoglobin concentration was 61 µg Hb/g feces (range 0-251 µg Hb/g) in men and 77 µg Hb/g feces (range 13-448 µg Hb/g) in women (p=0.76). In subjects with AN, men had a median fecal Hb concentration of 3.2 µg Hb/g feces (range 0-485 µg Hb/g) and women 2.6 µg Hb/g feces (IQR 0-670 µg Hb/g) (p=0.94).

Test characteristics

Performance characteristics of FIT for AN are provided in Table 2. Differences in test characteristics were not significant. At each of the pre-specified cut-offs, men had slightly higher positivity rates than women. The positivity rates ranged from 3.2% to 10.8% for the highest and lowest cut-off. The sensitivity for AN ranged from 40% (95% CI 29-52%) at FIT10 to 25% (95% CI 16-37%) at FIT40 in men, and from 35% (95% CI 24-49%) at FIT10 to 22% (95% CI 12-35%) at FIT40 in women. The specificity of FIT for AN tended to be lower in men when compared to women up to cut-offs of 20 µg Hb/g. The detection rate of AN was higher in men than women at all cut-offs. False positivity rates ranged from 1.1% (95% CI 0.5-2.3%) to 6.6% (95% CI 4.9-8.8%) for men and from 1.5% (95% CI 0.8-2.8%) to 5.5% (95% CI 4.0-7.6%) for women. True positivity rates ranged from 2.7% (95% CI 1.7-4.2%) to 4.2% (95% CI 2.9-6.1%) for men and from 1.8% (95% CI 1.0-3.2%) to 2.9% (95% CI 1.8-4.6%) for women. Sensitivity and specificity were calculated for the study population at multiple cut-offs ranging from 0 to 100 µg Hb/g feces (Figure 1). At an increasing cut-off, in both genders there is a relatively more rapid decline in sensitivity than an increase in specificity. Overall men had slightly higher sensitivities than women. For example, at a commonly used cut-off of 10 µg Hb/g feces women should have a lower cut-off to reach the same sensitivity and specificity as men.

Table 1 Findings at colonoscopy in men and women.

	Men n = 638	**Women** n = 618	**p-value**
Most advanced finding at colonoscopy *			
no histology	17 (2.7%)	3 (0.5%)	
no abnormalities	303 (47.5%)	357 (57.8%)	
SSA < 10mm	19 (3.0%)	17 (2.8%)	
HP	82 (12.8%)	77 (12.4%)	
TA < 10mm	149 (23.4%)	113 (18.3%)	
TA ≥ 10 mm	21 (3.3%)	16 (2.6%)	
SSA ≥ 10mm	1 (0.2%)	1 (0.2%)	
TVA	32 (5.0%)	25 (4.1%)	
VA	2 (0.3%)	0	
HGD	7 (1.1%)	6 (0.9%)	
CRC	5 (0.8%)	3 (0.5%)	
Total advanced neoplasia	68 (10.6%)	51 (8.3%)	0.15
Location of most advanced neoplasia* *			
distal / proximal (%)	51 (75) / 17 (25)	38 (75) / 13 (25)	0.95
Number of advanced neoplasia per participant n (%)**			0.55
1	53 (77.9%)	42 (82.4%)	
> 1	15 (22.1%)	9 (17.6)	

* no histology: removed polyp not retrieved for histology; SSA: sessile serrated adenoma; HP: hyperplastic polyp; TA: tubular adenoma; TVA: tubular villous adenoma; VA villous adenoma; HGD: high-grade dysplasia; CRC: colorectal cancer.

** only subjects with advanced neoplasia included (men n = 68 and women n = 51)

Figure 1 Sensitivity and specificity for men and women for all cut-offs ranging from 0 to 100 µg Hb/g feces.

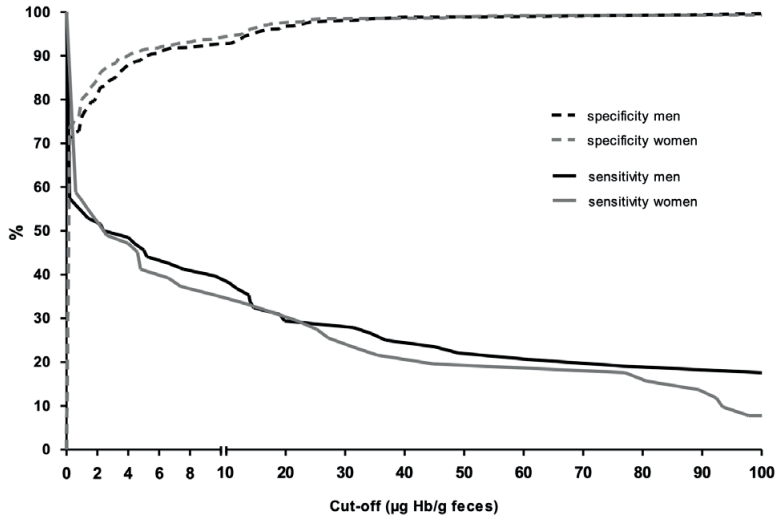


Table 2: Positivity rate, sensitivity, specificity, positive predictive value, negative predictive value, detection rate for men and women at different cut-offs.

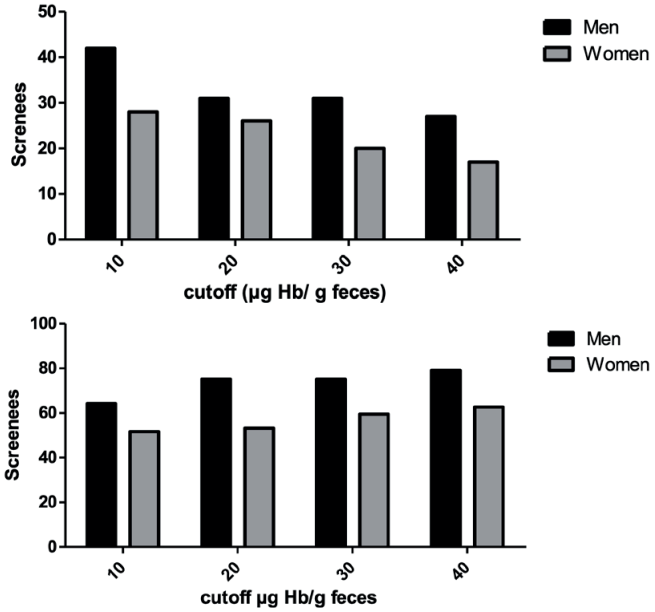
		Men		Women	
		%	(95%CI)	%	(95%CI)
FIT 10	PR	10.8	(8.6-13.5)	8.4	(6.5-10.9)
	sensitivity	39.7	(28.8-51.7)	35.3	(23.5-49.2)
	specificity	92.6	(90.2-94.5)	94.0	(91.7-95.7)
	PPV	39.1	(28.4-51.0)	34.6	(23.0-48.8)
	NPV	92.7	(90.4-94.7)	94.2	(91.9-95.8)
	DR	4.2	(2.9-6.1)	2.9	(1.8-4.6)
FIT 20	PR	6.3	(4.6-8.4)	5.0	(3.5-7.0)
	sensitivity	29.4	(19.8-41.2)	33.3	(21.8-47.2)
	specificity	96.4	(94.6-97.7)	97.5	(95.9-98.5)
	PPV	50.0	(35.0-65.0)	54.8	(37.4-71.1)
	NPV	92.0	(89.5-93.9)	94.2	(92.0-95.8)
	DR	3.1	(2.0-4.8)	2.8	(1.7-4.4)
FIT 30	sensitivity	29.4	(19.8-41.2)	25.5	(15.4-39.1)
	specificity	98.0	(96.3-98.8)	98.4	(97.0-99.2)
	PPV	62.5	(44.9-77.3)	59.0	(38.2-77.2)
	NPV	92.1	(89.6-94.0)	93.6	(91.4-95.3)
	DR	3.1	(2.9-4.8)	2.1	(1.2-3.6)
FIT 40	PR	3.8	(2.5-5.6)	3.2	(2.1-5.0)
	sensitivity	25.0	(16.1-36.6)	21.6	(12.4-34.9)
	specificity	98.8	(97.4-99.4)	98.4	(97.0-99.2)
	PPV	70.8	(50.2-85.4)	55.0	(33.6-74.7)
	NPV	91.7	(89.2-93.6)	93.3	(91.0-95.1)
	DR	2.7	(1.7-4.2)	1.8	(1.0-3.2)

FIT; fecal immunochemical test; FIT 10; cutoff level 10 µg/g feces; FIT 20; cutoff level 20 µg/g feces; FIT 30; cutoff level 30 µg/g feces; FIT 40; cutoff level 40 µg/g feces; PR: positivity rate; PPV positive predictive value; NPV: negative predictive value; DR: detection rate.

Detected lesions in absolute numbers

At all cut-offs, more lesions were detected as well as missed in men than in women (Figure 2 A, B). For all cut-offs the number needed to screen to identify one screenee with AN was higher in women than in men. It ranged from 38 to 56 subjects in women and from 24 to 38 subjects in men. Stepwise lowering cut-offs for men from FIT40 to respectively FIT30, FIT20 and FIT10 successively resulted in additional detection of 3, 0, and 7 AN. This required 8, 8, and 29 additional colonoscopies. Stepwise lowering the cut-offs for women from FIT40 to FIT10 successively resulted in additional detection of 2, 4 and 1 AN. This required 2, 4 and 21 additional colonoscopies.

Figure 2a) Detected advanced neoplasia per 1000 screenees (for 100% participation in absolute numbers). **b)** Missed advanced neoplasia per 1000 screenees (for 100% participation in absolute numbers).



DISCUSSION

In this colonoscopy-based screening program, we evaluated gender differences with respect to the efficacy of FIT screening in average risk individuals. Furthermore, we illustrated the effect of using different cut-offs on a broad spectrum of screening outcomes. Our study demonstrated that FIT had a higher sensitivity and lower specificity for AN in men than in women. By increasing the cut-off a relative more rapid decline in sensitivity was found than an increase in specificity for both genders. Furthermore, FIT had an overall higher PPV in men. When looking at diagnostic yield in absolute numbers, men had higher detection rates and miss rates of AN than women at all cut-offs. This last finding is of particular interest as in current literature little attention has been given to gender specific miss rates of lesions.

A strength of this study is that this cohort was set in a population-based screening setting, making these results representative for average-risk screening populations. Also, as all participants underwent both colonoscopy and FIT, it is a very suitable population to demonstrate actual differences and to estimate the number of missed lesions. However, to appreciate our findings some limitations need to be discussed. Firstly, this cohort consists of relatively small numbers and was not powered to detect differences in men and women. Another limitation is that only persons willing to undergo colonoscopy as primary

screening method participated, which could have led to a selection bias resulting in a population that is not representative of FIT-participants. Only 22% of all invitees decided to participate in colonoscopy screening, while FIT-screening generally has much higher participation rates around 60%.[19, 20] Nevertheless, the population includes only screening-naive average risk subjects, and therefore we think that the risk of selection-bias is limited.

The introduction of fecal immunochemical testing was an important step forward in population-based colorectal cancer screening. FIT enables simple, low burden primary screening at relatively low costs and has a high uptake. For this reason, an increasing number of countries have implemented FIT-based screening programs or are in the process of doing so.[21] This is mostly associated with a marked increase in colonoscopy demand. This asks for a strong focus on optimal use of limited resources.

Differences between men and women in terms of number of advanced lesions, location of lesions and fecal hemoglobin concentrations are becoming more evident. Dissimilarities in prevalence of AN between men and women have been well described with men having a substantial higher prevalence of AN than women.[10, 11] Consequently, research on tailored screening strategies become of significant importance. We are the first to describe the detection and miss rate of lesions in absolute numbers, showing that in men more lesions were both detected and missed for all cut-offs. This was especially the case at higher cut-offs. A previous Polish study showed that the number needed to screen in colonoscopy-based CRC screening to identify one screenee with advanced neoplasia was considerably higher in women than in men [10] Our data shows that these numbers also apply to colorectal cancer screening programs based on FIT. At each FIT cut-off, 14 to 18 more women needed to be screened to find one case of AN compared to men.

Differences in FIT screening between men and women can be explained by a combination of factors. It has been suggested that because men have a higher hemoglobin concentration in general, blood from bleeding polyps will contain more globin.[22] As FIT specifically detects globin in feces, blood from these polyps could be detected more frequently in men. This is supported by the fact that differences in fecal Hb concentration have been found in men and women.[15, 23] A second explanation could be that women have more right-sided lesions, as it is known that fecal occult blood testing may not be as sensitive for proximal lesions as it is for distal lesions.[14, 22, 24] Yet, our data did not show differences in location of AN between men and women. Another reason for gender differences in FIT test-characteristics, could be the differences in colonic transit time between men and women, with women having slower transit times.[25] A slower transit time could lead to more degradation of Hb and could decrease the likelihood of blood being detected by FIT.

An important question to be answered is how these results can be applied in colorectal cancer screening programs. Essentially, for gender-adjusted cut-offs in FIT-based CRC screening programs three scenarios are possible. These are the use of the same cut-off in both genders or, using a higher cut-off in men than women, or vice-versa. An increase of the cut-off for men compared to women can lead to a similar proportional sensitivity for detection of AN in both groups. As a consequence the difference in PPV between men and women would increase, with men having a substantial higher PPV. Also, a higher cut-off in men would lead to a further increase in miss rates of AN in men in absolute numbers and thus to a further increase in difference of miss rates in terms of absolute numbers of advanced lesions compared to women. Furthermore, using a lower cut-off for women would result in a higher rate of false-positive tests in women. The opposite strategy, i.e. increasing the cut-off for women compared to men, can lead to a similar miss rate in terms of absolute numbers, and to a similar PPV in both genders. It would however result in decreased sensitivity and detection rates for women. In this scenario a larger proportion of the colonoscopy capacity would be used for men. However, such a strategy could make sense given that men are at higher risk of AN and subsequently the development CRC.

Other gender-based CRC screening strategies besides adjusting the cut-off, include the use of different age ranges for screening, changing screening modality, or the use of different screening intervals. A German study showed that women reached equivalent levels of CRC-related mortality as men at a 4 to 8 years higher age [26]. Gender differences in other screening modalities, such as colonoscopy, sigmoidoscopy, fecal biomarkers and fecal DNA, have not yet been extensively investigated. However, using different methods or combinations of tests for men and women could optimize screening efficacy and should be further investigated.

With regard to gender differences in patient-education there is still much to gain. Information on miss rates of advanced lesions is an important issue in client information. At present, men and women are informed in the same manner about FIT-based CRC screening. These results helps to accurately inform the client about the gender-dependent risk of miss rates and detection of advanced lesions in a FIT based CRC screening program.

To conclude, colorectal cancer screening using FIT with the same cut-off for both genders results in a higher sensitivity and lower specificity for advanced neoplasia in men than in women. In absolute numbers more advanced neoplasia are detected and missed in men for all cut-offs. Following, tailored cut-off based on gender could either level sensitivity in men and women by using a lower cut-off in women, or level the amount of missed lesions when using a lower cut-off in men. Adjusting cut-offs based on gender can contribute to the efficacy of FIT-based CRC screening programs and optimize the use of available

endoscopy resources. In addition, individuals invited to attend a FIT based CRC screening should be informed accordingly about these gender differences.

Acknowledgements

Conceived idea for the study: E.J. Grobbee and M.C.W. Spaander; E.J. Grobbee, M.C.W. Spaander and E.J. Kuipers designed and conceptualized the study; Supervised execution of the study was done by M.C.W. Spaander; Responsible for data entry were E.M. Stoop, and T.R. de Wijkerslooth; Analysis and interpretation of data was done by E.J. Grobbee, E. Wieten, M.C.W. Spaander, E.J. Kuipers, I. Lansdorp-Vogelaar. The manuscript was drafted by E.J. Grobbee. E.Wieten, T.R. de Wijkerslooth, I. Lansdorp-Vogelaar, E.Dekker, P.M. Bossuyt, M.C.W. Spaander and E.J. Kuipers provided critical revision of the manuscript for important intellectual content.

REFERENCES

1. Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. Cancer Epidemiol Biomarkers Prev. 2009;18(6):1688-94.
2. Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer. 2010;116(3):544-73.
3. Shaukat A, Mongin SJ, Geisser MS, Lederle FA, Bond JH, Mandel JS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med. 2013;369(12):1106-14.
4. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy. 2013;45(1):51-9.
5. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100(7):1103-10.
6. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327-37.
7. 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.
8. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
9. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol. 2013;10(3):130-42.
10. Regula J, Rupinski M, Kraszewska E, Polkowski M, Pachlewski J, Orlowska J, et al. Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia. N Engl J Med. 2006;355(18):1863-72.
11. 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):1585-89.
12. Brenner H, Hoffmeister M, Birkner B, Stock C. Men with negative results of guaiac-based fecal occult blood test have higher prevalences of colorectal neoplasms than women with positive results. Int J Cancer. 2014;134(12):2927-34.
13. Ferlitsch M, Heinze G, Salzl P, Britto-Arias M, Waldmann E, Reinhart K, et al. Sex is a stronger predictor of colorectal adenoma and advanced adenoma than fecal occult blood test. Med Oncol. 2014;31(9):151.
14. Steele RJ, McClements P, Watling C, Libby G, Weller D, Brewster DH, et al. Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. Gut. 2012;61(4):576-81.
15. Kapidzic A, van der Meulen MP, Hol L, van Roon AH, Looman CW, Lansdorp-Vogelaar I, et al. Gender differences in fecal immunochemical test performance for early detection of colorectal neoplasia. Clin Gastroenterol Hepatol. 2015.

16. Ferlitsch M, Reinhart K, Pramhas S, Wiener C, Gal O, Bannert C, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA. 2011;306(12):1352-58.
17. de Wijkerslooth TR, de Haan MC, Stoop EM, Deutekom M, Fockens P, Bossuyt PM, et al. Study protocol: population screening for colorectal cancer by colonoscopy or CT colonography: a randomized controlled trial. BMC Gastroenterol. 2010;10:47.
18. de Wijkerslooth TR, Stoop EM, Bossuyt PM, Meijer GA, van Ballegooijen M, van Roon AH, et al. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia. Am J Gastroenterol. 2012;107(10):1570-8.
19. 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.
20. Kapidzic A, Grobbee EJ, Hol L, van Roon AH, van Vuuren AJ, Spijker W, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. Am J Gastroenterol. 2014;109(8):1257-64.
21. 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.
22. Brenner H, Haug U, Hundt S. Sex differences in performance of fecal occult blood testing. Am J Gastroenterol. 2010;105(11):2457-64.
23. Auge JM, Pellise M, Escudero JM, Hernandez C, Andreu M, Grau J, et al. Risk stratification for advanced colorectal neoplasia according to fecal hemoglobin concentration in a colorectal cancer screening program. Gastroenterology. 2014;147(3):628-36 e1.
24. Chacko L, Macaron C, Burke CA. Colorectal Cancer Screening and Prevention in Women. Dig Dis Sci. 2015;60(3):698-710.
25. Meier R, Beglinger C, Dederding JP, Meyer-Wyss B, Fumagalli M, Rowedder A, et al. Influence of age, gender, hormonal status and smoking habits on colonic transit time. Neurogastroenterol Motil. 1995;7(4):235-8.
26. Brenner H, Hoffmeister M, Arndt V, Haug U. Gender differences in colorectal cancer: implications for age at initiation of screening. Br J Cancer. 2007;96(5):828-31.

9

Accrediting for screening-related colonoscopy services: What is required of the endoscopist and of the endoscopy service?

E. Wieten, M.C.W. Spaander, E.J. Kuipers.

Best Practice & Research Clinical Gastroenterology, 2016

ABSTRACT

Colorectal cancer (CRC) screening is widely implemented to reduce CRC incidence and related mortality. The impact of screening as well as the balance between screening burden and benefits strongly depends on the quality of colonoscopy. Besides quality, safety of the endoscopic procedure and patient satisfaction are important outcome parameters for a screening program. Therefore the requirements for both CRC screening endoscopy services and endoscopists focus on technical aspects, patient safety, and patient experience. Stringent quality assurance by means of routine monitoring of quality indicators for the performance of endoscopists and endoscopy units is recommended. This allows setting minimum standards, targeted interventions, and enhancement of the overall quality of population screening. This reviews deals with guidelines and quality standards for colorectal cancer screening, with focus on both endoscopist and endoscopy services.

INTRODUCTION

In recent years, more than 50 countries have implemented organized or opportunistic population colorectal cancer (CRC) screening.[1] It has been convincingly demonstrated that CRC screening can reduce CRC-related mortality, as well as depending on the screening method, the incidence of the disease. Screening aims to lower the burden of cancer by discovering disease at an early, preclinical stage.[2-5] Population-based screening for CRC and precursor lesions can be effective provided that services and colonoscopies are of high quality.[6] Therefore, the European Union recommends to use evidence-based methods with quality assurance of the entire screening process.[7] To ensure that the experience is of high quality, safe and efficient, as well as people-oriented, services must take different domains of quality assurance into account. These are endoscopy/technical aspects, patients safety, and patients satisfaction.[8] The level of competency to perform high-quality endoscopy and to remove advanced lesions is not only dependent on the skills of the endoscopist, but also on the support team and the available facilities and equipment.[9] Screening enables known finite health gains, but also potential harms. Therefore, quality assurance of screening services and endoscopists is of utmost importance in CRC screening programs. This review describes the requirements for accrediting screening centers as well as individual endoscopists in a CRC screening program.

Organized versus opportunistic screening programs

The organization of CRC screening differs between countries.[1] In some countries, such as the United States, opportunistic programs have been in place for a long time, and cover a significant proportion of the population with proven effects on CRC incidence and mortality.[10] In most settings however, opportunistic programs are characterized by low or unknown participation rates, simultaneous frequent overuse of services by those subjects who do undergo screening, and lack of impact on national CRC incidence and mortality data. For these reasons, the European Union recommends organized screening programs.[11] In contrast to case-finding or opportunistic screening, organized programs provide a comprehensive data collection structure, which ensures evaluation and quality assessment. Centrally organized screening programs follow a predefined protocol, which enables systematic monitoring of the effectiveness of the program and process quality.[12] Also, potential harms can be surveyed, both at individual and systemic levels. If flaws in the screening program are identified, measures can be taken to improve and optimize the proposed screening.

Requirements for the endoscopy service

To ensure high-quality CRC screening programs, endoscopy services need to be efficient, safe, person-oriented and able to monitor key outcomes.[9, 13] In organized screening programs, accreditation of endoscopy services ensures that these conditions are met, in order to provide a minimum standard level of safety for participants.

The European guideline divides recommendations concerning endoscopy services into two categories. The first deals with planning and location of endoscopy services and the second with infrastructure and equipment.[9] Planning and location recommendations include that screening services be located in convenient locations for participants, that clinical services be accessible in a timely manner, and without compromising access to endoscopy services for symptomatic patients.[9] Infrastructure recommendations include proper facilities for pre-procedure assessment and post-procedure recovery with sufficient privacy to maintain dignity for the patient. The guideline also includes disinfection policies and procedures with the important remark that these should be compliant with national or international guidelines.

To perform high-quality endoscopy, remove advanced lesions and deal effectively with adverse events, a competent team and adequate equipment are required. Carbon dioxide insufflation is recommended for colonic endoscopic procedures since this has been proven to improve safety and reduce post-procedural discomfort. The available equipment needs to undergo regular safety checks and its use needs to be regularly trained. The risk of adverse events should be routinely assessed with every patient. In case of serious adverse events, which cannot be managed locally, the patient needs to be transferred safely for further care. The endoscopist and supporting staff must be competent to deliver high quality endoscopy. Patient comfort score and satisfaction should be monitored and periodically evaluated, followed by targeted measures where needed. [14] Examples of quality requirements for the screening endoscopy service and endoscopist are depicted in Table 1.

Requirements for the endoscopist

Colonoscopists are required to meet predetermined standards before involvement in a screening program, and to become subject to ongoing quality assurance. Several surrogate parameters have been developed to measure efficacy and professional quality on the level of the individual endoscopist. Commonly used parameters, which directly relate to detection of adenoma and cancer, are the quality of the bowel preparation, cecal intubation rate, colonoscopy withdrawal time, and adenoma detection rate (ADR).[15-17] In recent years, several studies have shown that adenoma detection rates directly relate to the risk of post-colonoscopy colorectal cancers (PCCRCs).[18, 19] In a Polish trial involving 45.026 individuals followed for an average four years after screening colonoscopy performed by 186 individual colonoscopists, 42 PCCRCs were diagnosed. [18] The hazard ratio for PCCRC was more than ten if the colonoscopy had been performed by an endoscopist with an ADR <20% compared to ≥20%. In a very similar trial from the US, 314.872 individuals underwent a colonoscopy performed by 136 gastroenterologists.[19] During follow-up up to 10 years, 712 patients were diagnosed with a PCCRC. The adjusted hazard ratio for interval cancer

Table 1 Examples of required quality indicators for screening endoscopy services and endoscopists with according minimum standards and targets.

Colonoscopy quality indicator	Examples	Minimum standard	Targets
Infrastructure			
	Proper facilities for pre-procedure assessment and post-procedure recovery.	Demonstrable	According to (inter)national guidelines
	Contamination policies and procedures.	Demonstrable	According to (inter)national guidelines
	Adequate equipment for high quality colonoscopies.	Safety checks and regular training of the available and new equipment	Documentation on safety checks and regular training performed
	Percentage of colonoscopies where CO_2 insufflation is used.	100% use of CO2	–
	Capable to record, monitor and deliver quality and auditable indicators [9], [13]	Demonstrable	Full registration of all quality indicators in 100% of the colonoscopies performed
Performance			
Documentation of quality standards	Percentage of colonoscopies with adequate bowel preparation [8], [28], [29]	≥90%	≥95%
	Percentage of colonoscopies with complete visualization of the cecum [9]	≥90%	≥97%
	Percentages of colonoscopies where at least one adenoma is found (ADR) [8], [18], [19]	≥30–35%	≥40%
	Percentages of negative colonoscopies with a withdrawal time ≥6 min [17]	90%	95%
Documentation of auditable outcomes	Percentage of patients with discomfort [14]	100% monitoring of comfort scores per every colonoscopy	–
	Number of colonoscopies performed per endoscopist per year [9]	≥300 per year	–
	Recording colonoscopic findings, procedural methods and interventions [49]	MAP/MAP+/Polyp retrieval ≥90%	Polyp retrieval ≥95%
Documentation of auditable outcomes	Percentage of patients with discomfort [14]	100% monitoring of comfort scores per every colonoscopy	–
Fundamental outcomes			
	Recording of adverse events [9], [13]	–	Post polypectomy bleeding rate <1 per 100
		–	Perforation rate <1 per 1000
	Registration of interval cancers	–	–

was 0.52 for patients of endoscopists with adenoma detection rates in the highest quintile compared to those in the lowest quintile. Every stepwise 1% increase of ADR reduced the risk for PCCRC with 3%.[19]

These and similar trials strongly emphasized the need for quality assurance in screening colonoscopy. This is further enhanced by the fact that the procedure is associated with risks of serious complications.[20]

Quality standards are measurable outcomes for which there is an evidence base that supports a minimum standard. [8] They drive quality to higher standards while setting limits to identify suboptimal performance. There is a range of accepted quality standards for the individual endoscopist performing screening colonoscopy.

Bowel preparation

Adequate colon cleansing is mandatory to the endoscopic visualization of the colonic mucosa. The quality of bowel preparation thus influences endoscopy efficacy, and directly affects both cecal intubation rates and polyp detection rates.[21-24] In addition, inadequate bowel preparation can result in longer procedure times, increased screenee burden, additional costs, shorter surveillance intervals, re-scheduled procedures, and alternative diagnostic investigations [25]. A recent prospective study showed that bowel preparation together with detection of advanced neoplasia, were the only two independent significant predictors of the need for a second-look colonoscopy in a CRC screening program.[25] In this study involving 1250 screenees with a positive faecal immunochemial test, 8.6% underwent a second-look colonoscopy within 1 year, of which 13% was due to poor bowel preparation at the initial investigation. This finding highlights the impact that poor colon cleansing can have on colonoscopy resources. Finally, patient's (dis)comfort and (in) convenience with regards to bowel preparation may have an effect on attendance rates of endoscopy in CRC screening programs. A randomized study showed that the proportion of individuals complaining of serious adverse events from the bowel preparation was higher (odds ratio 5.17) in those undergoing a colonoscopy than in individuals undergoing flexible sigmoidoscopy.[26] In this study screenees prepared for sigmoidoscopy by means of a self-administered enema (133 ml, 22% sodium phosphate) at home two hours before the procedure; for colonoscopy, an oral preparation with sodium phosphate solution (2L), starting in the afternoon preceding the scheduled appointment, was used. The most common complaints were abdominal distension and pain.

It is recommended that at least 85% of patients referred for colonoscopy and at least 90% of those who do so for screening purposes should present with an adequately cleansed colon.[8, 27, 28] In order to achieve adequate colon cleansing, the United States Multi-Society

Task Force (USMSTF) and the European Society of Gastrointestinal Endoscopy (ESGE) both recommend split-dose bowel preparation regimens. [27, 28] A split-dose regime implies that the bowel cleansing dose is given in two portions either on the day before and the day of the exam, or a same day split-dose regimen in the case of afternoon procedures. Preferably, the intake of the second dose should start 4–6 hours before the procedure time, and is completed two hours before the procedure.[27] The efficacy of split-dose versus day-before bowel cleansing regimens were assessed in more than 40 randomized trials. Two meta-analyses concluded that split-dose regimens lead to a significantly higher proportion of patients with adequately cleansed bowels compared with day-before regimens (odds ratios 2.5 – 3.7).[29, 30] As a result, split-dose compared to day-before regimes lead to higher adenoma detection rates (ADR) and potentially higher detection rates of flat polyps.[17, 30, 31]. Split dosing also significantly decreased the proportion of patients that complained of nausea (odds ratio 0.55) or who were unable to complete the bowel preparation (odds ratio 0.53).[32] This increases the willingness to repeat the preparation and exam when needed (odds ratio 1.8).[28]

Complications need to be taken into account, when prescribing bowel cleansing medication. The ESGE advises against the routine use of sodium phosphate for bowel preparation because of safety concerns. Sodium phosphate compounds can lead to significant fluid and electrolyte shifts and are thus particularly contra-indicated in patients with hypertension, those taking diuretics or renin-angiotensin blockers, or who suffer from heart failure, chronic kidney disease, or liver failure.[33] In these patients, PEG is the only recommended bowel preparation.[28] Screening units should record and report on the bowel cleansing regimen used, and the proportion of screenees with adequately cleansed bowels.

Cecal intubation rates

Cecal intubation is defined as passage of the colonoscope tip to a point proximal to the ileocecal valve so that the entire cecal cap is visible. Cecal intubation rates (CIRs) reflect the proportion of patients in whom the entire colon has been reached with the colonoscope.[34, 35] It is recommended that CIRs be reported as unadjusted rates, i.e. rates that are not corrected for incompleteness due to luminal obstruction or poor bowel preparation. Unadjusted rates are preferred because the corrections could be applied differently and adjusted rates can therefore be difficult to compare.

When combined with measures of bowel cleansing and withdrawal time, CIR serves as a measure of the completeness of inspection of the colonic surface. Cecal intubation rates of 96% and higher have been reported in CRC screening programs.[35-39], hence the

target of ≥95% (unadjusted) cecal intubation rate is required in colonoscopy screening.[9] In symptomatic patients a slightly lower target level of ≥90% is generally used.[34]

Complete colonoscopy to the cecum is sometimes difficult and cannot be performed due to patient discomfort, looping, colonic redundancy, severe diverticulosis, or adhesions. Adjunct endoscopic techniques such as water immersion, abdominal pressure, patient positioning and use of magnetic endoscopic imaging can overcome many of these challenges. [40] Screening units and endoscopists should record CIRs, and report the reason for not reaching the cecum.

Adenoma detection rates

Multiple studies have shown that polyp and adenoma detection rates (ADR) vary between endoscopists.[41-45] ADR is defined as the number of colonoscopies in which at least one adenomatous lesion is identified divided by total amount of colonoscopies.[46] The variation in ADR between endoscopists and the demonstration that ADRs inversely correlate with the risk for PCCRC have become the rationale for the creation of targets for adenoma detection.[18, 19, 34, 43]. Interval cancers are defined by the World Endoscopy Organization as cancer diagnosed after a screening test or examination in which no cancer is detected and before the date of the next recommended examination.[47] Two large studies formed the basis for the current ADR target recommendation, being 30% in men and 20% in women in an average risk population aged ≥ 50 years undergoing primary screening colonoscopy.[18, 19] These targets should be set higher in screening programs that are based on faecal occult blood testing as primary tests, followed by colonoscopy after a positive fecal test. In the UK screening program with guaiac faecal occult blood testing, minimal adenoma detection rates are set at 35%, with a preferred target ADR >40%.[8]

Since adenoma detection rates do not take the number of adenomas per positive screenee into account, and the presence of an advanced adenoma is the most important predictor for adenoma recurrence, some programs also assess the mean number of adenomas per procedure and the mean number of adenomas per positive procedure (respectively MAP and MAP+).[48] These measures are thought to provide additional information about colonoscopy performance in addition to ADR, and have also been shown to differ between endoscopists.[8]

Colonoscopy withdrawal time

In 2006, a first study showed an association between colonoscopy withdrawal time (WT) and detection of advanced neoplasia.[42] Colonoscopists with a mean WT of more than 6 minutes had 2.5 times higher detection rates of advanced neoplasia compared to colonoscopists with WT less than 6 minutes. This finding has been confirmed in other

studies. A multicenter randomized trial assessing the association between multiple parameters and ADR in a colorectal cancer screening program, showed that WT was the only factor related to ADR.[17] In another recent Canadian cohort study involving more than 18,000 asymptomatic individuals, WT and CIR were associated with the detection of screen-relevant lesions (defined as CRC, serrated and adenomatous polyps).[16] In this study, both an average WT over 6 minutes and a >20% detection rate of screen-relevant lesions were associated with an almost 6- to 8-fold decreased risk of PCCRCs.

Colonoscopy WT is thus an important parameter to measure when performing a screening colonoscopy. The ESGE guidelines recommend a mean WT ≥6 minutes in negative colonoscopies performed in average risk subjects for CRC screening.[49]

Endoscopy volumes

In contrast to colonoscopies in symptomatic patients, screening endoscopies involve healthy individuals. This further emphasizes the need for attention to the potential benefits to risks ratio. Several studies have shown that low numbers of examinations per endoscopist are associated with an increased risk of complications. [20, 50] In a large retrospective study from Canada, patients had a 3-fold higher odds of bleeding or perforation if their colonoscopy was performed by endoscopists with the lowest annual volume of examinations compared to patients who had their colonoscopy performed by endoscopists with the highest annual volume.[20] This risk of complications was increased below a threshold of 300 colonoscopies per year. As a result, current European guidelines recommend that each endoscopist participating in a colorectal cancer screening program should perform at least 300 procedures per year to maintain adequate competence.[9] This also ensures sufficient sample size to reliably assess quality indicators as mentioned above.

Screening accreditation and strategies for suboptimal performance

Accreditation of screening endoscopists

To outline the service and quality indicators expected and in order to ensure that a high standard of service is provided, (organized) screening programs provide accreditation processes for screening endoscopists and endoscopy units. These accreditation processes lead to a certificate of competency to perform screening colonoscopy. There is a considerable disparity between countries in accreditation practice for screening endoscopists. All have different criteria for accreditation that vary in the methods of assessment, minimum number of procedures, and key quality criteria.[51] The responsibility for accreditation for colonoscopy also differs between countries. Responsibility for accreditation for colonoscopy can be the responsibility of the national gastroenterology society such as in Canada, or of the national screening program such as in the Netherlands. In the USA, training in gastrointestinal endoscopy is outlined in the Gastroenterology

Core Curriculum, while England has one of the most rigorous accreditation processes. Here, a Joint Advisory Group for GI Endoscopy (JAG) manages the Screening Assessor Accreditation System (SAAS) process with defined parameters for endoscopy training and national endoscopy standards [52, 53].

Accreditation of endoscopy services

The Global Rating Scale (GRS) is a quality assurance program that was developed in England to assess patient-centered care in endoscopy.[54] The development of the GRS was prompted by the introduction of the national colorectal cancer screening program and by shortcomings in the quality of endoscopy.[55] The GRS is suitable to assess patient experiences, and can be used as a benchmark tool for quality and safety in endoscopy departments. More countries are of have implemented the UK Global Rating Scale in an effort to improve and maintain high standards of endoscopy service.[56]

Besides measuring quality indicators, it is also important to consider how suboptimal performances ought to be managed. A systematic review showed that interventions targeting endoscopist performance have generally been ineffective for improving adenoma or polyp detection rates.[57] However, studies included in this systematic review had small sample sizes and a lack of randomized designs. Recently, a large trial randomized endoscopy screening centers with suboptimal performance (with the centre leader having an ADR ≤25%) to either a Train-Colonoscopy-Leaders programme (assessment, hands-on training, post-training feedback) or feedback only (individual quality measures). It showed that leadership training led to a significant 7.1% higher increase in ADR compared to feedback only.[58] hese kinds of solutions to increase endoscopist performance will help to optimize CRC screening programs by decreasing the variation in quality between screening centers, and increase colonoscopy quality without decreasing colonoscopy resources by excluding centers that did not reach quality benchmarks.

SUMMARY

In recent years, more than 50 countries worldwide have implemented CRC screening programs. Colorectal cancer screening can lead to significant population health gains, but can also harm healthy subjects. High quality colonoscopies are required in order to achieve CRC screening efficacy and reduce CRC-related mortality. Centrally organized screening and accreditation of screening services are recommended by the European Union and provide the possibility of systematic monitoring and quality assessment. An important parameter at the endoscopist's level is the ADR. Multiple studies have shown that ADR differs between endoscopists, and recent evidence convincingly showed an inverse correlation between ADR and risk of PCCRC, emphasizing the importance of high quality endoscopies. Future research is needed to address current practice in the identification

of early endoscopically resectable CRCs and endoscopic resection techniques used to remove them accurately in a CRC screening setting in order to improve screening program, increase quality and prevent unnecessary surgery.

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.
2. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev. 2007(1):CD001216.
3. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med. 1993;329(27):1977-81.
4. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375(9726):1624-33.
5. Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med. 2012;366(8):687-96.
6. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy. 2013;45(1):51-9.
7. Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) Off J Eur Union. 2003:34–38.
8. Lee TJ, Rutter MD, Blanks RG, Moss SM, Goddard AF, Chilton A, et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. Gut. 2012;61(7):1050-7.
9. Valori R, Rey JF, Atkin WS, Bretthauer M, Senore C, Hoff G, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Quality assurance in endoscopy in colorectal cancer screening and diagnosis. Endoscopy. 2012;44 Suppl 3:SE88-105.
10. Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, Noone AM, et al. Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer. 2012;118(9):2338-66.
11. Malila N, Senore C, Armaroli P, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Organisation. Endoscopy. 2012;44 Suppl 3:SE31-48.
12. Senore C, Inadomi J, Segnan N, Bellisario C, Hassan C. Optimising colorectal cancer screening acceptance: a review. Gut. 2015;64(7):1158-77.
13. Sung JJ, Ng SC, Chan FK, Chiu HM, Kim HS, Matsuda T, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. Gut. 2015;64(1):121-32.
14. Ekkelenkamp VE, Dowler K, Valori RM, Dunckley P. Patient comfort and quality in colonoscopy. World J Gastroenterol. 2013;19(15):2355-61.
15. Rex DK, Petrini JL, Baron TH, Chak A, Cohen J, Deal SE, et al. Quality indicators for colonoscopy. Am J Gastroenterol. 2006;101(4):873-85.
16. Hilsden RJ, Dube C, Heitman SJ, Bridges R, McGregor SE, Rostom A. The association of colonoscopy quality indicators with the detection of screen-relevant lesions, adverse events, and postcolonoscopy cancers in an asymptomatic Canadian colorectal cancer screening population. Gastrointest Endosc. 2015;82(5):887-94.

17. Jover R, Zapater P, Polania E, Bujanda L, Lanas A, Hermo JA, et al. Modifiable endoscopic factors that influence the adenoma detection rate in colorectal cancer screening colonoscopies. Gastrointest Endosc. 2013;77(3):381-9 e1.
18. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362(19):1795-803.
19. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370(14):1298-306.
20. Rabeneck L, Paszat LF, Hilsden RJ, Saskin R, Leddin D, Grunfeld E, et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. Gastroenterology. 2008;135(6):1899-906, 906 e1.
21. Froehlich F, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc. 2005;61(3):378-84.
22. Harewood GC, Sharma VK, de Garmo P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. Gastrointest Endosc. 2003;58(1):76-9.
23. Hsu CM, Lin WP, Su MY, Chiu CT, Ho YP, Chen PC. Factors that influence cecal intubation rate during colonoscopy in deeply sedated patients. J Gastroenterol Hepatol. 2012;27(1):76-80.
24. Lebwohl B, Kastrinos F, Glick M, Rosenbaum AJ, Wang T, Neugut AI. The impact of suboptimal bowel preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. Gastrointest Endosc. 2011;73(6):1207-14.
25. Grobbee EJ, Kapidzic A, van Vuuren AJ, van Leerdam M, Lansdorp-Vogelaar I, Looman CW, et al. Second-Look Colonoscopies and the Impact on Capacity in FIT-Based Colorectal Cancer Screening. Am J Gastroenterol. 2015;110(7):1072-7.
26. Senore C, Ederle A, Fantin A, Andreoni B, Bisanti L, Grazzini G, et al. Acceptability and side-effects of colonoscopy and sigmoidoscopy in a screening setting. J Med Screen. 2011;18(3):128-34.
27. Hassan C, Bretthauer M, Kaminski MF, Polkowski M, Rembacken B, Saunders B, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) guideline. Endoscopy. 2013;45(2):142-50.
28. Johnson DA, Barkun AN, Cohen LB, Dominitz JA, Kaltenbach T, Martel M, et al. Optimizing adequacy of bowel cleansing for colonoscopy: recommendations from the US multi-society task force on colorectal cancer. Gastroenterology. 2014;147(4):903-24.
29. Martel M, Barkun AN, Menard C, Restellini S, Kherad O, Vanasse A. Split-Dose Preparations Are Superior to Day-Before Bowel Cleansing Regimens: A Meta-analysis. Gastroenterology. 2015;149(1):79-88.
30. Gurudu SR, Ramirez FC, Harrison ME, Leighton JA, Crowell MD. Increased adenoma detection rate with system-wide implementation of a split-dose preparation for colonoscopy. Gastrointest Endosc. 2012;76(3):603-8 e1.
31. Parra-Blanco A, Nicolas-Perez D, Gimeno-Garcia A, Grosso B, Jimenez A, Ortega J, et al. The timing of bowel preparation before colonoscopy determines the quality of cleansing, and is a significant factor contributing to the detection of flat lesions: a randomized study. World J Gastroenterol. 2006;12(38):6161-6.
32. Kilgore TW, Abdinoor AA, Szary NM, Schowengerdt SW, Yust JB, Choudhary A, et al. Bowel preparation with split-dose polyethylene glycol before colonoscopy: a meta-analysis of randomized controlled trials. Gastrointest Endosc. 2011;73(6):1240-5.
33. Connor A, Tolan D, Hughes S, Carr N, Tomson C. Consensus guidelines for the safe prescription and administration of oral bowel-cleansing agents. Gut. 2012;61(11):1525-32.

34. Rex DK, Schoenfeld PS, Cohen J, Pike IM, Adler DG, Fennerty MB, et al. Quality indicators for colonoscopy. Am J Gastroenterol. 2015;110(1):72-90.
35. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. N Engl J Med. 2000;343(3):169-74.
36. Leung JW, Mann SK, Siao-Salera R, Ransibrahmanakul K, Lim B, Cabrera H, et al. A randomized, controlled comparison of warm water infusion in lieu of air insufflation versus air insufflation for aiding colonoscopy insertion in sedated patients undergoing colorectal cancer screening and surveillance. Gastrointest Endosc. 2009;70(3):505-10.
37. Cadoni S, Gallittu P, Sanna S, Fanari V, Porcedda ML, Erriu M, et al. A two-center randomized controlled trial of water-aided colonoscopy versus air insufflation colonoscopy. Endoscopy. 2014;46(3):212-8.
38. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G, et al. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. New Engl J Med. 2000;343(3):162-8.
39. Schoenfeld P, Cash B, Flood A, Dobhan R, Eastone J, Coyle W, et al. Colonoscopic screening of average-risk women for colorectal neoplasia. N Engl J Med. 2005;352(20):2061-8.
40. Villa NA, Pannala R, Pasha SF, Leighton JA. Alternatives to Incomplete Colonoscopy. Curr Gastroenterol Rep. 2015;17(11):43.
41. Imperiale TF, Glowinski EA, Juliar BE, Azzouz F, Ransohoff DF. Variation in polyp detection rates at screening colonoscopy. Gastrointest Endosc. 2009;69(7):1288-95.
42. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med. 2006;355(24):2533-41.
43. Chen SC, Rex DK. Endoscopist can be more powerful than age and male gender in predicting adenoma detection at colonoscopy. Am J Gastroenterol. 2007;102(4):856-61.
44. Hixson LJ, Fennerty MB, Sampliner RE, McGee D, Garewal H. Prospective study of the frequency and size distribution of polyps missed by colonoscopy. J Natl Cancer Inst. 1990;82(22):1769-72.
45. Bressler B, Paszat LF, Vinden C, Li C, He J, Rabeneck L. Colonoscopic miss rates for right-sided colon cancer: a population-based analysis. Gastroenterology. 2004;127(2):452-6.
46. Rex DK, Petrini JL, Baron TH, Chak A, Cohen J, Deal SE, et al. Quality indicators for colonoscopy. Gastrointest Endosc. 2006;63(4 Suppl):S16-28.
47. Sanduleanu S, le Clercq CM, Dekker E, Meijer GA, Rabeneck L, Rutter MD, et al. Definition and taxonomy of interval colorectal cancers: a proposal for standardising nomenclature. Gut. 2015;64(8):1257-67.
48. van Heijningen EM, Lansdorp-Vogelaar I, Kuipers EJ, Dekker E, Lesterhuis W, Ter Borg F, et al. Features of adenoma and colonoscopy associated with recurrent colorectal neoplasia based on a large community-based study. Gastroenterology. 2013;144(7):1410-8.
49. Rembacken B, Hassan C, Riemann JF, Chilton A, Rutter M, Dumonceau JM, et al. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). Endoscopy. 2012;44(10):957-68.
50. Singh H, Penfold RB, DeCoster C, Kaita L, Proulx C, Taylor G, et al. Colonoscopy and its complications across a Canadian regional health authority. Gastrointest Endosc. 2009;69(3 Pt 2):665-71.
51. Kelly NM, Moorehead J, Tham T. Is the 'driving test' a robust quality indicator of colonoscopy performance? World J Gastrointest Endosc. 2010;2(4):112-20.
52. Schiller LR. The core curriculum in gastroenterology: training the gastroenterologists of the future. Am J Gastroenterol. 2007;102(5):919-20.
53. www.thejag.org.uk.
54. Global Rating Scale UK. Available from http://wwwgrsnhsuk. 01.05.2010.

55. Bowles CJ, Leicester R, Romaya C, Swarbrick E, Williams CB, Epstein O. A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow? Gut. 2004;53(2):277-83.
56. Sint Nicolaas J, de Jonge V, Korfage IJ, Ter Borg F, Brouwer JT, Cahen DL, et al. Benchmarking patient experiences in colonoscopy using the Global Rating Scale. Endoscopy. 2012;44(5):462-72.
57. Corley DA, Jensen CD, Marks AR. Can we improve adenoma detection rates? A systematic review of intervention studies. Gastrointest Endosc. 2011;74(3):656-65.
58. Kaminski MF, Anderson J, Valori R, Kraszewska E, Rupinski M, Pachlewski J, et al. Leadership training to improve adenoma detection rate in screening colonoscopy: a randomised trial. Gut. 2015;65(4):616-24.

10

General discussion and future perspectives

SUMMARY

The aims of this thesis were to compare different fecal occult blood tests for colorectal cancer screening and to explore tailored FIT-based screening strategies. We assessed the incidence of interval cancers after a negative guaiac fecal occult blood test and a negative fecal immunochemical test. Second, two different fecal immunochemical test assays were compared with regard to detection of advanced neoplasia, participation rate, and ease of use. Next, tailored screening strategies were explored using fecal immunochemical tests. Finally, requirements for accrediting screening centers as well as individual endoscopists in a colorectal cancer screening program were described.

Colorectal cancer (CRC) is particularly suitable for screening and, in recent years, more CRC screening has been implemented worldwide.[1-3] Currently offered CRC screening programs can be divided in either invasive screening methods, such as colonoscopy or flexible sigmoidoscopy, or non-invasive methods, including fecal occult blood testing (FOBT).

Fecal occult blood testing for colorectal cancer screening

Two types of fecal occult blood testing have been widely available for CRC screening: guaiac fecal occult blood testing (gFOBT) and fecal immunochemical testing (FIT). gFOBT has been the most commonly used stool screening method for years. However, gFOBT has been rapidly replaced by FIT, as FIT has been shown to be more sensitive for the detection of CRC as well as its precursors than gFOBT. FIT also allows for single stool testing, is easier to handle, is associated with higher uptake and may provide quantitative test results.[4-8] The latter characteristic enables to adjust the positivity cut-off to match available resources. However, data on the incidence rate of test-related interval cancers were limited. And, although interval cancer rate is considered a key quality indicator in screening programs, no data were available on how gFOBT and FIT compare with regard to these interval cancers incidence rates. In **Chapter 2** we performed a systematic literature search and meta-analysis to determine and compare the incidence rates of test-related interval colorectal cancers of gFOBT and FIT in population-based CRC screening programs. We found that the incidence of interval cancer after a negative gFOBT was higher than after a negative FIT. We also found that for every FIT interval cancer, 2.6 colorectal cancers were detected; for gFOBT the ratio between interval and detected colorectal cancers was 1:1.2. Although high heterogeneity was shown among studies included in this meta-analysis, and although included FIT studies were generally based on low positivity cut-offs, the study results favor the use of FIT over gFOBT as screening test for CRC. The outcomes further help to adequately inform screenees about the risk of interval cancers after a negative fecal occult blood test.

From January 2014 onwards, a nationwide FIT-based CRC screening program has been gradually rolled out in the Netherlands using the FOB-Gold (Sentinel, Italy). However, previous research in The Netherlands was done using the OC-Sensor (Eiken, Japan). Head-to-head comparisons with enough power to determine if these tests were equivalent in detecting advanced neoplasia were not available. We conducted two large prospective population-based studies within the Dutch CRC screening program to compare these two FIT assays on multiple outcomes. In **Chapter 3**, we demonstrated equivalence between the two FIT assays in detection of advanced neoplasia in a large prospective paired accuracy study at a positivity cut-off level of ≥15 µg Hb/g feces. The paired study also serves as an example for how to assess and possibly improve screening effectiveness within an ongoing program. In **Chapter 4**, we compared participation rates and ease

of use of FOB-Gold and OC-sensor within the Dutch CRC screening program. Screening invitees were asked to complete both OC-sensor and FOB-Gold and assess ease of use and preference of FIT brand by questionnaire. In parallel, we compared participation rates in a randomized trial and assessed the proportions of non-analyzable tests. This study found small, but statistically significant, differences in ease of use in favor of FOB-Gold. Most participants did not express a preference for either FIT. Those that did, preferred FOB-Gold over OC sensor. Despite these differences, our randomized trial showed that participation in the Dutch population based CRC screening program was not influenced by the type of FIT offered. Furthermore, non-analyzable test proportions were small (0.1% versus 0.4%), and we found no differences in non-analyzability between the tests. As these two large prospective population-based studies show that detection rates of advanced neoplasia are equivalent and participation rates similar, other features may guide decision-making for selecting a FIT in a CRC screening program.

We also estimated how these two FIT assays compare in detecting advanced neoplasia at different test positivity cut-offs in **Chapter 5**. FIT assays vary in analytical performance due to a range of factors, including anti-heme antibody characteristics and assay optimization, buffer composition and volume and sample tube design. These differences influence the measured fecal hemoglobin concentration, FIT positivity rate, error rates, and capacity to detect AN.[9, 10] The two widely used FITs had significantly different distributions of reported hemoglobin concentration and yielded different positivity rates at equal thresholds. However, they performed similarly in detecting advanced neoplasia at a similar positivity rate. When comparing and implementing these FIT assays in a screening program, the desired positivity rate that identifies participants to be referred for colonoscopy should first be set, guided by available resources and feasibility.

Tailored colorectal cancer screening strategies

Next, we explored tailored CRC screening strategies with FIT. Tailored screening strategies that we explored included two-sample FIT screening, the use of different positivity cut-offs, and tailored screening based on age categories and gender.

In **Chapter 6**, we assessed the number of advanced neoplasia detected by two-sample screening with FIT assays taken from the same bowel movement. We found that the proportion of participants with discordant FIT results almost equaled the proportion of those with two positive tests, while a generally low positivity cut-off was used. Given the high rate of advanced neoplasia detection in this group with two discordant test results (27%), colonoscopy should strongly be considered in case a screenee has one positive FIT. The findings did however not particularly favor two-sample over one-sample FIT

screening. When screening strategies are considered that improve advanced neoplasia detection rates, drop in specificity should be taken into account.

When facing limited colonoscopy resources, various FIT-based screening strategies are optional. We calculated if increasing the positivity cut-off and screening age affects colonoscopy yield, missed lesions, and colonoscopy demand in **Chapter 7**. In a population-based CRC screening cohort, we found that FIT positivity rates, detection rates and the positive predictive value to detect advanced neoplasia all increase with age. Both increasing the screening starting age and increasing the positivity cut-off resulted in a substantial reduction in colonoscopy demand, at the cost of a similar number of advanced neoplasia missed.

We illustrated the effects of gender-tailored screening in **Chapter 8** and assessed the effects on miss rates of advanced neoplasia. We showed that in absolute numbers more advanced neoplasia are detected and missed in men than in women at all positivity cut-offs. Gender based positivity cut-offs could level sensitivity for both men and women by using a lower cut-off in women, or level the amount of lesions missed when using a lower cut-off in men.

Depending on the desired aim of CRC screening policy-makers, tailored screening strategies are optional. When colonoscopy resources are limited, tailored screening based on age categories, gender and FIT positivity cut-off are all considerable options.

Screening endoscopists and the endoscopy service

Population-based screening for CRC and precursor lesions can be effective provided that services and colonoscopy are of high quality.[11] Therefore, the European Union recommends to use evidence-based methods with quality assurance of the entire screening process.[12] Screening enables known finite health gains, but also potential harms. Therefore, quality assurance of screening services and endoscopists is of utmost importance. The review of **Chapter 9** describes the requirements for accrediting screening centers as well as individual endoscopists in a CRC screening program.

Future Perspectives

Worldwide, CRC screening programs have been implemented over the past years.[3] However, improvements regarding the quality and accuracy of CRC screening programs and optimization of screening strategies can still be made. Besides, the optimal screening strategy may not be similar for different demographic areas, due to differences in the screening population, insurance systems and access to health care systems. Also, these aspects may change over time. Therefore, continuous evaluation of new strategies

to optimize screening program effectiveness and reduce possible harms is crucial to guarantee screening quality, and should be performed on a routine basis.[13]

In Europe, the quantitative immunochemical tests are recommended as test of choice for population CRC screening.[6] However, more than 50 FIT brands are widely available.[14] In this thesis, we compared two FIT brands and their ability to detect advanced neoplasia, participation rates, and ease of use. Although literature is available on other FIT brands, more research is required. Also, many other features should be compared including cost-effectivity, ease of use for laboratory staff or other stakeholders involved in FIT analysis, suitability for transport, keeping quality of the tubes, analyzer features, capacity, speed, analytical performance, sample stability, easy of handling, safety during postage and labeling. Depending on context and setting, more studies are warranted to evaluate these other aspects of FIT.

One should realize that CRC screening using FIT is far from optimal. A meta-analysis showed that, at generally used test positivity cut-offs, the pooled sensitivity and specificity of FIT for CRC were 0.79 and 0.94, respectively. The perfect screening test has a sensitivity and specificity of 1.0. In addition, we found in our meta-analysis that for every CRC missed by FIT, 2.6 CRCs were detected. Potential new screening targets include CRC-specific proteins, fecal deoxyribonucleic acid (DNA) or detection by odor material.[11, 15, 16] However, none of these alternative screening tests have proven to be superior to FIT or direct colonoscopy so far.

An important feature in screening is uptake of the test. A screening test or series of tests must be acceptable for the population.[1] As others have stated, "the best screening test, is the test that gets done, and gets done well".[17] Still much can be improved with regard to reaching high-risk groups for CRC screening. A recent study even argued that greater benefit, at lower cost, could be achieved by increasing participation rates for unscreened older and higher-risk persons than lowering the screening starting age to 45 years.[18]

Finally, future research can focus on surveillance strategies for patients after polyp removal. Current guidelines recommend frequent surveillance colonoscopies for patients after polyp removal, which contributes to a significant burden of colonoscopy demand.[19-22] Currently, there is a lack in evidence of clinical trials on optimal surveillance colonoscopy strategies. The upcoming European Polyp Surveillance (EPoS) study results may contribute to more evidence-based knowledge on this topic.[23]

REFERENCES

1. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease] Principios y metodos del examen colectivo para identificar enfermedades. Bol Oficina Sanit Panam. 1968;65(4):281-393.
2. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med. 1993;329(27):1977-81.
3. 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.
4. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100(7):1103-10.
5. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327-37.
6. Halloran SP, Launoy G, Zappa M, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:SE65-87.
7. Kuipers EJ, Spaander MC. Colorectal Cancer Screening by Colonoscopy, CT-Colonography, or Fecal Immunochemical Test. J Natl Cancer Inst. 2016;108(2):djv383.
8. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol. 2013;10(3):130-42.
9. 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. 2016;66(11):1975-82.
10. Chiang TH, Chuang SL, Chen SL, Chiu HM, Yen AM, Chiu SY, et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cutoff concentration in a nationwide colorectal cancer screening program. Gastroenterology. 2014;147(6):1317-26.
11. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy. 2013;45(1):51-9.
12. Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) Off J Eur Union. 2003:34–38.
13. Bell KJ, Bossuyt P, Glasziou P, Irwig L. Assessment of changes to screening programmes: why randomisation is important. BMJ. 2015;350:h1566.
14. Daly JM, Xu Y, Levy BT. Which Fecal Immunochemical Test Should I Choose? J Prim Care Community Health. 2017;8(4):264-77.
15. Bosch LJW, de Wit M, Pham TV, Coupe VMH, Hiemstra AC, Piersma SR, et al. Novel Stool-Based Protein Biomarkers for Improved Colorectal Cancer Screening: A Case-Control Study. Ann Intern Med. 2017;167(12):855-66.
16. Sonoda H, Kohnoe S, Yamazato T, Satoh Y, Morizono G, Shikata K, et al. Colorectal cancer screening with odour material by canine scent detection. Gut. 2011;60(6):814-9.
17. Allison JE. The best screening test for colorectal cancer is the one that gets done well. Gastrointest Endosc. 2010;71(2):342-5.

18. Ladabaum U, Mannalithara A, Meester RGS, Gupta S, Schoen RE. Cost-Effectiveness and National Effects of Initiating Colorectal Cancer Screening for Average-Risk Persons at Age 45 Years Instead of 50 Years. Gastroenterology. 2019;157(1):137-48.
19. Atkin WS, Valori R, Kuipers EJ, Hoff G, Senore C, Segnan N, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. Endoscopy. 2012;44 Suppl 3:SE151-63.
20. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR, et al. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology. 2012;143(3):844-57.
21. Hassan C, Quintero E, Dumonceau JM, Regula J, Brandao C, Chaussade S, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy. 2013;45(10):842-51.
22. Grobbee EJ, Kapidzic A, van Vuuren AJ, van Leerdam M, Lansdorp-Vogelaar I, Looman CW, et al. Second-Look Colonoscopies and the Impact on Capacity in FIT-Based Colorectal Cancer Screening. Am J Gastroenterol. 2015;110(7):1072-7.
23. Jover R, Bretthauer M, Dekker E, Holme O, Kaminski MF, Loberg M, et al. Rationale and design of the European Polyp Surveillance (EPoS) trials. Endoscopy. 2016;48(6):571-8.

A

Dutch summary
Contributing authors
Bibliography
PhD portfolio

INTRODUCTIE

Darmkanker is een veelvoorkomende ziekte. Wereldwijd is darmkanker de derde meest voorkomende kanker in mannen en tweede in vrouwen.[1] De kans om gedurende het leven darmkanker te krijgen ligt rond de 5%. De behandeling van darmkanker is over de laatste jaren sterk verbeterd, maar nog steeds overlijdt 40-50% van de patiënten met symptomatische darmkanker uiteindelijk aan gemetastaseerde ziekte.[2] Darmkanker ontwikkelt zich meestal uit een adenoom, de voorloper van darmkanker.[3] Echter, niet alle adenomen ontwikkelen zich uiteindelijk tot darmkanker. De ontwikkeling van een vroeg adenoom tot invasieve darmkanker duurt jaren.[4]

Via een bevolkingsonderzoek kunnen mensen met darmkanker, of voorlopers daarvan, opgespoord worden vóórdat de ziekte zich heeft openbaart. De vroege detectie van darmkanker en zijn voorlopers maakt de kans op genezing groter.[5] Er zijn verschillende methodes beschikbaar voor een bevolkingsonderzoek naar darmkanker. Grofweg zijn er twee methodes, invasieve en niet-invasieve onderzoeken. Bij de invasieve onderzoeken wordt de dikke darm in beeld gebracht middels een darmonderzoek. Een voorbeeld van niet-invasieve onderzoeken is de fecaal occult bloed test (FOBT). De FOBT meet of er bloed in de ontlasting zit. In het geval dat er bloed wordt gevonden, wordt de deelnemer verwezen om een darmonderzoek te ondergaan.

Momenteel zijn er twee type occult bloed testen wereldwijd beschikbaar, de guaiac fecaal occult bloed test (gFOBT) en de fecaal immunochemische test (FIT). Gerandomiseerde en gecontroleerde onderzoeken hebben laten zien dat een bevolkingsonderzoek met gFOBT is geassocieerd met 15%-33% daling in darmkanker-gerelateerde mortaliteit.[6-8] Wereldwijd wordt momenteel het gebruik van de gFOBT voor bevolkingsonderzoeken naar darmkanker vervangen door gebruik van FIT, omdat de FIT meer sensitief bleek te zijn voor de detectie van darmkanker en voorlopers van darmkanker ten opzichte van de gFOBT.[9,10] FIT detecteert humaan-specifiek globine van bloed, terwijl gFOBT reageert met heem, waaronder heem van geconsumeerd vlees. Andere voordelen van de FIT ten opzichte van de gFOBT zijn dat de test makkelijker af te nemen is, geassocieerd is met hogere deelname en kwantitatief van aard is.[11,12] Deze laatste eigenschap zorgt ervoor dat de gehanteerde afkapwaarde kan worden aangepast aan de beschikbare capaciteit van dikke darmonderzoeken.[11,12]

Fecaal occult bloed testen voor een bevolkingsonderzoek darmkanker

De graad van interval kankers is een zeer belangrijke kwaliteit indicator van een bevolkingsonderzoek.

Intervalkankers worden gedefinieerd als kankers die ontstaan na een negatieve test en voordat de eerst volgende test zou moeten plaatsvinden.[13] Er was nog weinig onderzoek gedaan naar de incidentie van intervalkankers na een negatieve gFOBT en FIT. In **hoofdstuk 2** voerden we een systematisch literatuur onderzoek uit en een meta-analyse om de incidentie van deze interval kankers te berekenen. We vonden dat de incidentie van interval kankers na een negatieve gFOBT hoger was dan na een negatieve FIT. We vonden ook dat voor elke FIT interval kanker, 2.6 darmkankers werden opgespoord en voor gFOBT was deze verhouding interval kankers ten opzichte van opgespoorde kankers 1:1.2. Deze resultaten ondersteunen het gebruik van FIT ten opzichte van gFOBT. Daarnaast helpen de resultaten om de screening populatie te informeren over het risico op deze interval kankers.

In Nederland is het bevolkingsonderzoek darmkanker gestart in 2014. Individuen tussen de 55 en 75 jaar oud worden elke 2 jaar uitgenodigd om een FIT, de FOB-Gold (Sentinel, Italy), uit te voeren en in geval van een verhoogde uitslag worden zij uitgenodigd voor een darmonderzoek. Echter, eerdere proefbevolkingsonderzoeken waren gebaseerd op een ander merk FIT, de OC-Sensor (Eiken, Japan). We hebben daarom een groot wetenschappelijk onderzoek gedaan om deze twee testen met elkaar te vergelijken in hun vermogen om darmkanker en voorlopers van darmkanker, te samen gevorderde poliepen of advanced neoplasie genoemd, op te sporen. We nodigden meer dan 40.000 mensen uit om deze twee testen uit te voeren in dezelfde ontlasting. In **hoofdstuk 3** lieten we zien dat beide testen equivalent waren in de opsporing van advanced neoplasie bij een gehanteerde afkapwaarde van de test van ≥15 µg Hb/gram feces. In **hoofdstuk 4** vergeleken we dezelfde testen met elkaar, maar dan op gebruikersgemak van de test en deelname percentage. We vonden kleine, maar significante, verschillen in gebruikersgemak ten voordele van de FOB-Gold. De meeste deelnemers hadden geen voorkeur voor één van beide testen.

We vergeleken de beide testen ook in hun vermogen om advanced neoplasie op te sporen bij verschillende afkapwaarden in **hoofdstuk 5**. De FOB-Gold en OC-Sensor detecteerden significant verschillende hemoglobine concentraties in de ontlasting met verschillende bijbehorende percentages van positieve testen. Echter, bij eenzelfde positiviteitsgraad, was hun opsporend vermogen van advanced neoplasie wel gelijk. Dus wanneer deze testen worden geïmplementeerd of met elkaar worden vergeleken, moet eerst het gewenste percentage positieve testen worden vastgesteld.

Darmkanker screening strategieën op maat

Factoren die gebruikt kunnen worden om screening op maat aan te bieden zijn onder andere geslacht, leeftijd, leefstijl, omgevingsfactoren, genen, aantal aangeboden testen

en de gehanteerde afkapwaarde van de test. In **hoofdstuk 6** bekeken we het diagnostisch vermogen van een bevolkingsonderzoek met twee testen en bekeken we hoe kan worden omgegaan met een deelnemer die twee testresultaten heeft. We vonden dat de proportie deelnemers met twee verschillende testuitslagen (één positieve en één negatieve uitslag) bijna gelijk was aan de proportie deelnemers met twee positieve uitslagen. In de groep met twee verschillende testuitslagen was het percentage met gevorderde poliepen bovendien 27%. Daarom moet een darmonderzoek sterk worden aangeraden aan mensen met twee verschillende FIT-uitslagen.

In **hoofdstuk 7** analyseerden we de percentages positieve testen en detectiegraad van gevorderde poliepen van FIT over verschillende leeftijdscategorieën en berekende de impact op de vraag naar darmonderzoeken bij gebruik van verschillende afkapwaarden van de test. We vonden dat de percentages positieve testen, de detectiegraad en de positief voorspellende waarde om gevorderde poliepen te detecteren, allemaal toenamen met de leeftijd. Zowel het verhogen van de start leeftijd met screenen als het verhogen van de afkapwaarde van de test resulteerde in een lagere vraag naar darmonderzoeken, ten koste van een gelijk percentage gemiste gevorderde poliepen.

In **hoofdstuk 8** onderzochten we het effect van screening op maat naar geslacht. We stelden vast dat in absolute getallen meer gevorderde poliepen worden gedetecteerd en gemist bij mannen dan bij vrouwen bij alle afkapwaarden. Dientengevolge, kan een gelijke sensitiviteit van gevorderde poliepen voor mannen en vrouwen nagestreefd worden, door een lagere afkapwaarde voor vrouwen dan voor mannen te hanteren. Anderzijds kan een gelijk aantal gemiste gevorderde poliepen worden nagestreefd voor mannen en vrouwen door een lagere afkapwaarde bij mannen te hanteren.

Screening ewndoscopisten en endoscopie centra

De impact van screening naar darmkanker en de balans tussen voor- en nadelen van screening hangt nauw samen met de kwaliteit van het darmonderzoek. Naast kwaliteit, zijn tevens veiligheid van het darmonderzoek en patiënt tevredenheid belangrijke parameters in een bevolkingsonderzoek. In **hoofdstuk 9** beschrijven we de benodigdheden en voorwaarden waar endoscopie centra en endoscopisten aan moeten voldoen voor een bevolkingsonderzoek darmkanker.

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer* 2019; 144(8): 1941-53.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; 127(12): 2893-917.
3. Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. Natural history of untreated colonic polyps. *Gastroenterology* 1987; 93(5): 1009-13.
4. 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): 1585-89.
5. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nat Rev Dis Primers* 2015; 1: 15065.
6. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 1996; 348(9040): 1472-77.
7. Kronborg O, Fenger C, Olsen J, Jørgensen OD. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996; 348(9040): 1467-71.
8. 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): 1365-71.
9. Hol L, Wilschut JA, van Ballegooijen M, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. *Br J Cancer* 2009; 100(7): 1103-10.
10. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut* 2015; 64(8): 1327-37.
11. Kuipers EJ, Spaander MC. Colorectal Cancer Screening by Colonoscopy, CT-Colonography, or Fecal Immunochemical Test. *J Natl Cancer Inst* 2016; 108(2): djv383.
12. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. *Nat Rev Clin Oncol* 2013; 10(3): 130-42.
13. Sanduleanu S, le Clercq CM, Dekker E, et al. Definition and taxonomy of interval colorectal cancers: a proposal for standardising nomenclature. *Gut* 2015; 64(8): 1257-67.

CONTRIBUTING AUTHORS

Patrick M. Bossuyt
Department of Clinical Epidemiology, Biostatistics and Bioinformatics
Academic Medical Center Amsterdam
Amsterdam, The Netherlands

Wichor M. Bramer
Medical Library
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Marco J. Bruno
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Evelien Dekker
Department of Gastroenterology and Hepatology
Academic Medical Center Amsterdam
Amsterdam, The Netherlands

Esmée J. Grobbee
Department of Gastroenterology and Hepatology
Academic Medical Center Amsterdam
Amsterdam, The Netherlands

Bettina E. Hansen
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Clasine M. de Klerk
Department of Gastroenterology and Hepatology
Academic Medical Center Amsterdam
Amsterdam, The Netherlands

Ernst J. Kuipers
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Iris Lansdorp-Vogelaar
Department of Public Health
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Daan Nieboer
Department of Public Health
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Stella A. Nieuwenburg
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Chris R. Ramakers
Department of Clinical Chemistry
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Eline H. Schreuders
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Manon C.W. Spaander
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Annemieke van der Steen
Regional Organization for Population Screening
South-West Netherlands
Rotterdam, The Netherlands

Esther M. Stoop
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

Thomas R. de Wijkerslooth
Department of Gastroenterology and Hepatology
Academic Medical Center Amsterdam
Amsterdam, The Netherlands

BIBLIOGRAPHY

Wieten E, de Klerk CM, Lansdorp-Vogelaar I, Bossuyt PM, Dekker E, Spaander MCW. A quarter of participants with advanced neoplasia have discordant results from 2-sample fecal immunochemical tests for colorectal cancer screening. Clinical Gastroenterology and Hepatology (in press).

de Klerk CM, **Wieten E**, van der Steen A, Ramakers CR, Kuipers EJ, Hansen BE, Lansdorp-Vogelaar I, Bossuyt PM, Spaander MCW, Dekker E. Participation and ease of use in colorectal cancer screening: a comparison of two fecal immunochemical tests. American Journal of Gastroenterology. 2019 Mar;114(3):511-518.

de Klerk CM, **Wieten E**, Lansdorp-Vogelaar I, Bossuyt PM, Spaander MCW, Dekker E. Performance of two faecal immunochemical tests in detecting advanced neoplasia at different positivity thresholds: a cross-sectional study of the Dutch national colorectal cancer screening programme**.** The Lancet Gastroenterology & Hepatology. 2019 Feb;4(2):111-118.

Wieten E, de Klerk CM, van der Steen A, Ramakers CR, Kuipers EJ, Hansen BE, Lansdorp-Vogelaar I, Bossuyt PM, Dekker E, Spaander MCW. Equivalent accuracy of 2 quantitative fecal immunochemical tests in detecting advanced neoplasia in an organized colorectal cancer screening program. Gastroenterology. 2018 Nov;155(5):1392-1399.e5.

Wieten E, Schreuders EH, Grobbee EJ, Nieboer D, Bramer WM, Lansdorp-Vogelaar I, Bruno MJ, Kuipers EJ, Spaander MCW. Incidence of faecal occult blood test interval cancers in population-based colorectal cancer screening: a systematic review and meta-analysis. Gut. 2019 May;68(5):873-881.

Schreuders EH, **Wieten E**, Kuipers EJ, Spaander MCW. Optimizing Fecal Immunochemical Testing For Colorectal Cancer Screening. Clinical Gastroenterology and Hepatology. 2017 Oct;15(10):1498-1499.

Grobbee EJ, **Wieten E**, Hansen BE, Stoop EM, Wijkerslooth TR, Lansdorp-Vogelaar I, Bossuyt PM, Dekker E, Kuipers EJ, Spaander MCW. Fecal immunochemical test-based colorectal cancer screening: the gender dilemma. United European Gastroenterology Journal. 2017 apr;5(3):448-454.

Wieten E, Schreuders EH, Nieuwenburg SAV, Hansen BE, Lansdorp-Vogelaar I, Kuipers EJ, Bruno MJ, Spaander MCW. Effects of Increasing Screening Age and Fecal Hemoglobin Cutoff Concentrations in a Colorectal Cancer Screening Program. Clinical Gastroenterology and Hepatology. 2016 Dec;14(12):1771-1777.

Wieten E, Spaander MCW, Kuipers EJ. Accrediting for screening-related colonoscopy services: what is required of the endoscopist and of the endoscopy service? Best Practice & Research Clinical Gastroenterology. 2016 Jun;30(3):487-495.

Wieten E, Brouwer JT, Quispel R, Veldt BJ. Letter: scoring models in alcoholic hepatitis. Alimentary Pharmacology & Therapeutics. 2015 Jul;42(1):126.

Wieten E, van der Linden-Schrever BE, Sonneveld E, Veerman AJ, Pieters R. CD73 (5'-nucleotidase) expression has no prognostic value in children with acute lymphoblastic leukemia. Leukemia. 2011 Aug;25(8):1374-1376.

PHD PORTFOLIO

Name PhD student:	Els Wieten
PhD period:	June 2014 – December 2019
Erasmus MC department:	Gastroenterology and Hepatology
Promotor:	prof. dr. M.J. Bruno
Co-promotor:	dr. M.C.W. Spaander

Courses and workshops	Year	Workload
Endnote workshop. Erasmus MC, Rotterdam	2014	6 hours
Systematic literature research Pubmed workshop. Erasmus MC, Rotterdam	2014	6 hours
Systematic literature research other databases workshop. Erasmus MC, Rotterdam	2014	6 hours
Biostatistical Methods I: Basic Principles Methodology. NIHES, Erasmus MC, Rotterdam	2014	5.7 ECTs
Survival Analysis Course. MolMed, Erasmus MC, Rotterdam	2014	12 hours
Integrity in scientific research. Department of medical ethics and philosophy, Erasmus MC, Rotterdam	2015	9 hours
Basiscursus Regelgeving Klinisch Onderzoek. Consultatiecentrum patiëntgebonden onderzoek (CPO), Erasmus MC, Rotterdam *Certificate Good Clinical Practice obtained*	2015	24 hours
Biomedical English Writing and Communication. Erasmus MC, Rotterdam	2015	40 hours

Oral presentations	Year	Workload
Long-term follow-up of patients hospitalized for alcoholic hepatitis. Najaarscongres, Digestive Disease Days, Veldhoven, The Netherlands.	2014	28 hours
Endoscopically resectable colorectal cancer in a FIT-based screening population in the Netherlands: progress still to be made. Najaarscongres, Digestive Disease Days, Veldhoven, The Netherlands.	2016	28 hours
FIT comparison study as an example of a research project along the national screening program: what can be learned from a study in progress? MedOCC autumn meeting, Amsterdam, The Netherlands.	2016	16 hours
Time for all countries to get FIT! Why and how? Symposium 'New ways in cancer management: for quality of life'. Hamburg, Germany.	2017	28 hours
Differences in colonoscopy associated costs between primary colonoscopy and colonoscopy after positive FIT in colorectal cancer screening. Voorjaarscongres, Digestive Disease Days, Veldhoven, The Netherlands.	2017	16 hours
Pooled incidence of fecal occult blood test interval cancers in colorectal cancer screening; a systematic review and meta-analysis. Voorjaarscongres, Digestive Disease Days, Veldhoven, The Netherlands.	2017	16 hours
Meta-analysis of gFOBT and FIT interval cancers. World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting, Chicago, United States of America.	2017	16 hours
Comparison of two brands of fecal immunochemical tests within the Dutch nationwide CRC screening program. Voorjaarscongres, Digestive Disease Days, Veldhoven, The Netherlands.	2018	16 hours
Comparison of diagnostic yield of two brands of fecal immunochemical tests within the Dutch nationwide colorectal cancer screening program. Digestive Disease Week, Washington D.C., United States of America.	2018	16 hours

Poster presentations	Year	Workload
Positive predictive value increases with age in a FIT-based colorectal screening program. Digestive Disease Week, Washington D.C., United States of America.	2015	16 hours
Positive predictive value increases with age in a FIT-based colorectal screening program. United European Gastroenterology Week, Barcelona, Spain.	2015	16 hours
Endoscopic removal of early stage colorectal cancer in a FIT-based screening population: progress still to be made. Digestive Disease Week, San Diego, United States of America.	2016	16 hours
Endoscopic removal of early stage colorectal cancer in a FIT-based screening population: progress still to be made. United European Gastroenterology Week, Vienna, Austria.	2016	16 hours
Differences in colonoscopy associated costs between primary colonoscopy and colonoscopy after positive FIT in colorectal cancer screening. Digestive Disease Week, Chicago, United States of America.	2017	16 hours
Pooled incidence of fecal occult blood test interval cancers in colorectal cancer screening; a systematic review and meta-analysis. Digestive Disease Week, Chicago, United States of America.	2017	16 hours
Incidence of fecal occult blood test interval cancers in colorectal cancer screening; a systematic review and meta-analysis. United European Gastroenterology Week, Barcelona, Spain.	2017	16 hours

(Inter)national conferences attended	Year	Workload
Digestive Disease Days. Veldhoven, The Netherlands	2014	16 hours
Voorjaarscongres, Digestive Disease Days. Veldhoven, The Netherlands	2015	16 hours
Digestive Disease Week. Washington D.C., United States of America	2015	28 hours
United European Gastroenterology Week. Barcelona, Spain	2015	28 hours
Digestive Disease Week. San Diego, United States of America	2016	28 hours
Najaarscongres. Digestive Disease Days. Veldhoven, The Netherlands.	2016	16 hours
Voorjaarscongres, Digestive Disease Days. Veldhoven, The Netherlands.	2017	16 hours
Digestive Disease Week. Chicago, United States of America	2017	28 hours
United European Gastroenterology Week. Barcelona, Spain	2017	28 hours
Voorjaarscongres, Digestive Disease Days. Veldhoven, The Netherlands.	2018	16 hours
Digestive Disease Week. Washington D.C., United States of America	2018	28 hours

Seminars attended	Year	Workload
6e Lagerhuisdebat Hepatitis B and C, Utrecht, The Netherlands	2014	6 hours
10e Jaarsymposium Gastro-enterologie. Rotterdam, The Netherlands.	2015	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. Washington D.C., United States of America	2015	8 hours
30th Erasmus Liver Day. Rotterdam, The Netherlands	2015	8 hours
Wetenschapsmiddag arts-assistenten vereniging (AAV). Erasmus MC, Rotterdam	2015	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. Barcelona, Spain.	2015	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. San Diego, United States of America	2016	8 hours
Regionale gastro-enterologie bijscholing Zuid-West (pancreas). Rotterdam, The Netherlands.	2016	6 hours
Symposium 'New ways in cancer management: for quality of life'. Hamburg, Germany.	2017	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. Chicago, United States of America *Co-chair of Expert Working Group Meeting 'Right-Sided Lesions and Interval Cancers'*	2017	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. Barcelona, Spain.	2017	8 hours
World Endoscopy Organisation (WEO) Colorectal Cancer Screening Meeting. Washington D.C., United States of America	2018	8 hours

Awards

Oral free paper prize (2017)

Certificate of recognition for scientific accomplishment as an early career investigator (2018)

Memberships

Nederlands Vereniging voor Gastro-enterologie (NVGE)

Nederlands Vereniging voor Hepatologie (NVH)

Reviewing activities

BMC Gastroenterology

Clinical Gastroenterology and Hepatology

European Journal of Epidemiology

Gut

International Journal of Cancer