OPTIMIZING COLORECTAL CANCER SCREENING

new strategies and insights

Eline Schreuders

Stellingen behorende bij het proefschrift getiteld

OPTIMIZING COLORECTAL CANCER SCREENING: NEW STRATEGIES AND INSIGHTS

1. Fecal immunochemical tests are superior to guaiac fecal occult blood tests in detecting colorectal cancer in average-risk individuals. *(This thesis)*
2. The variety of cut-off concentrations and strategies used for fecal immunochemical test screening, illustrates that a one-size-FITs-all approach does not exist. *(This thesis)*
3. Although the detection rates of advanced neoplasia by fecal immunochemical testing differ substantially among age groups, age partitioned cut-off concentrations are not recommended. *(This thesis)*
4. Fecal hemoglobin concentration below the cut-off is an independent predictor of incident advanced neoplasia. *(This thesis)*
5. Repeated two-sample fecal immunochemical test screening does not increase the yield of advanced neoplasia compared to one-sample screening. *(This thesis)*
6. In today's world "just Google it" is considered to be the answer to everything, but for health related questions or translations this should be treated with caution. *(This thesis, and Patil et al. BMJ 2014)*
7. A positive screening result does not increase participants' level of anxiety or depression, nor decrease participants' level of health-related quality of life. *(Kirkøen et al. British Journal of Cancer 2016)*
8. Nederlandse huisartsen moeten meer weet hebben van de testeigenschappen van fecaal occult bloedtesten en zich bewust zijn van de juiste indicatiestelling. *(Klein-Puite et al. Huisarts Wet 2015)*
9. Het beoordelingsproces van nieuwe geneesmiddelen zoals momenteel geaccepteerd door registratie autoriteiten staat in scherp contrast met de huidige standaarden voor systematische reviews en meta-analyses. *(Geneesmiddelenbulletin 2016; 50: 65)*
10. To make yourself feel happier, you should help others. *(Borgonovi et al. Soc Sci Med. 2008)*
11. "Pas als je de moed toont je eigen weg te gaan, toont de weg zich aan jou". *(Paulo Coelho)*

Eline Schreuders
Rotterdam, 15 december 2016

OPTIMIZING COLORECTAL CANCER SCREENING

new strategies and insights

Eline Schreuders

The printing of this thesis has been financially supported by the Department of Gastroenterology and Hepatology, Erasmus MC University Medical Center, Rotterdam; Nederlandse Vereniging voor Gastroenterologie; Bevolkingsonderzoek Zuid-West; Chipsoft B.V.; CorporateHealth International; DaklaPack Clinical Trials B.V.; Eiken Chemical Europe; Erbe Nederland B.V.; Ferring B.V.; Norgine B.V.; Olympus Nederland B.V.; Pentax Medical B.V.; Tramedico B.V.; Zambon Nederland B.V.

ISBN:	978-94-028-0411-9
Printing:	Ipskamp Printing, Enschede, the Netherlands
Cover and layout:	Lyanne Tonk, Persoonlijkproefschrift.nl

Optimizing Colorectal Cancer Screening
new strategies and insights

Optimaliseren van darmkanker screening
nieuwe strategieën en inzichten

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 15 december 2016 om 09.30 uur

door

Elisabeth Helena Schreuders

geboren te Utrecht

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor:	Prof.dr. E.J. Kuipers
Overige leden:	Prof.dr. H.J. de Koning Prof.dr. S.J.O. Veldhuyzen van Zanten Prof.dr. E. Dekker
Copromotor:	Dr. M.C.W. Spaander

Paranimfen:	E.J. Grobbee E. Wieten

CONTENTS

PART ONE

Introduction

CHAPTER 1.1

General introduction

Colorectal cancer (CRC) is the third leading cause of cancer-related mortality.[1] The general aim of screening is to filter an average-risk population to detect and treat those at an asymptomatic and early stage in order to reduce disease burden. CRC is an ideal target for population screening. Besides aiming to improve outcomes through earlier diagnosis, detection and treatment of "pre-cancers" can prevent the development of CRC. Various test methods and strategies are available for CRC screening. Because a screening program focuses on an in principle healthy population, it necessitates quality assurance for every aspect of the program.

This thesis is divided into four parts. Part I introduces the main topics of the thesis. These chapters are followed by the aims and outline of the thesis. Part II and III encompass the main body of the thesis. The last part (IV) summarizes and discusses the main findings of the thesis and provides directions for future research.

CHAPTER 1.2

Colorectal cancer screening; a global overview of existing programs

E.H. Schreuders
A. Ruco
L. Rabeneck
R.E. Schoen
J.J.Y. Sung
G.P. Young
E.J. Kuipers

Gut. 2015; 64 (10), 1637-1649

SUMMARY

Colorectal cancer (CRC) ranks third among the most commonly diagnosed cancers worldwide, with wide geographical variation in incidence and mortality across the world. Despite proof that screening can decrease CRC incidence and mortality, CRC screening is only offered to a small proportion of the target population worldwide. Throughout the world there are widespread differences in CRC screening implementation status and strategy. Differences can be attributed to geographical variation in CRC incidence, economic resources, health care structure, and infrastructure to support screening such as the ability to identify the target population at risk and cancer registry availability. This review highlights issues to consider when implementing a CRC screening program and gives a worldwide overview of CRC burden and the current status of screening programs, with focus on international differences.

KEY MESSAGES

- A successful screening program for a major disease like colorectal cancer (CRC) requires comprehensive collaboration among multiple parties for an optimal effect in terms of gain in life-years, quality of life and cost-efficiency.
- Despite well-developed CRC screening guidelines, implementation of screening is markedly different among countries and regions in the world.
- In addition to CRC incidence, the impact of the disease relative to other health problems and the capacity to treat should be taken into account when developing a CRC screening strategy.
- The discrepancy between CRC incidence and the offer of organized or opportunistic screening remains an ongoing concern.
- Screening measures and quality indicators of screening should be reported, allowing national evaluation and international comparison to improve CRC screening quality.

BACKGROUND

Worldwide colorectal cancer (CRC) is the third most common cancer in men (746 000 cases, 10.0% of the total) and the second in women (614 000 cases, 9.2% of the total).[2] There is however wide geographical variation in CRC incidence and mortality, with very similar patterns in men and women. The age-standardized incidence rates (ASRi) vary ten-fold in both sexes worldwide. When comparing world regions as classified by the United Nations, the highest estimated rates occur in the Australia/New Zealand region (ASRi 44.8 and 32.2 per 100 000 in men and women respectively), and the lowest in Western Africa (4.5 and 3.8 per 100 000). Almost 55% of CRC cases occur in more developed regions.[2] However, in many developing countries (including some parts of Africa), there is the possibility of under-reporting because cancer registries are lacking or have incomplete coverage. Large disparities exist between high-income and low-income countries in the proportion of their populations covered by cancer registries.[3]

The lifetime risk of developing CRC in many regions is around 5%. Approximately 45% of persons diagnosed with CRC die as a result of the disease, despite treatment.[4] Treatment modalities have largely improved over the past decade. Treatment has modestly improved disease outcome and extended survival in patients with advanced and metastatic disease. But, these advancements have been accompanied by markedly increased treatment costs. As a result, modelling studies have shown that various screening strategies are cost-saving.[5] Most CRCs develop from a preclinical precursor, the adenoma. The progression from early adenoma to invasive cancer takes years.[6 7] The high incidence, long preclinical phase, recognizable and treatable precursor, the high cost of treatment, and the correlation of mortality with disease stage make CRC highly suitable for population screening.[8-10] This has been confirmed by randomized controlled trials (RCTs) that have formed the basis for international guidelines recommending CRC screening.[11-14] Despite these recommendations, screening is currently only offered to a small proportion of the target population.

The goal of this review is to address various aspects to consider for implementing a successful screening program and to give an overview of screening programs worldwide, with a focus on international differences.

METHODS

For this review, national and international guidelines on CRC screening were evaluated. We collected information on CRC screening program characteristics from guidelines, through national governmental websites and international contact persons including public health researchers, those responsible for the development and implementation of screening programs, and participants of the Colorectal Cancer Screening Committee of the World Endoscopy Organization.[15] A literature search in PUBMED and The Cochrane Central Register of Controlled Trials was performed using the following keywords: colorectal cancer (CRC) screening, guidelines, Europe, America, Canada, Asia, Australia, New Zealand, RCTs, colonoscopy, guaiac fecal occult blood test (gFOBT), fecal immunochemical test for hemoglobin (FIT), flexible sigmoidoscopy (FS), CT colonography (CTC), DNA-marker and video capsule endoscopy. To evaluate and compare screening programs, we used the universally applicable CRC screening indicators established by the International Colorectal Cancer Screening Network (ICRCSN) based on the criteria of the International Agency for Research on Cancer (IARC).[16 17] To report screening indicators regarding FIT screening, the Fecal Immunochemical TesTs for hemoglobin Evaluation Reporting (FITTER) guidelines were followed.[18] We used the recommended reporting units of microgram hemoglobin per gram feces (µg Hb/g) rather than nanogram hemoglobin per milliliter buffer (ng Hb/ml) to ensure comparability of results.[19] The term 'average risk population' used in this review refers to an asymptomatic population who is at average risk for CRC. The age range of this population is influenced by national guidelines and varies per study but is mainly over age 50 and constantly over age 40 years. For the overview of current status of screening programs, at least the top 10 countries with highest age-standardized incidence rates (ASRi) for each world continent were included.

SCREENING METHODS

Detection and removal of cancer precursors can reduce CRC incidence and mortality. Early detection of CRC allows less invasive treatment, with lower morbidity, mortality, and treatment cost. The implementation of a CRC screening program requires that strategic decisions be addressed. One is the selection of a screening modality, which can be a non-invasive or an invasive test.

Non-invasive stool tests

Non-invasive stool tests include guaiac fecal occult blood tests (gFOBT) and fecal immunochemical test for hemoglobin (FIT). These inexpensive tests detect microscopic amounts of blood by targeting either heme (gFOBTs) or human globin (FITs). A meta-analysis of four RCTs concluded that annual or biennial gFOBT screening had no effect on CRC incidence (in 3 out of the 4 studies included in the analysis) but led to an average 16% reduction in CRC-related mortality.[20] The impact of the gFOBT is limited by the poor to moderate sensitivity for advanced adenomas and cancer (Table 1.2.1).[21] For this reason, gFOBTs are typically used on multiple bowel movements per screening, and are implemented in repeated screening rounds. In contrast, FITs have a higher sensitivity for both adenoma and cancer even with a single sample per screening round (Table 1.2.1). Moreover, unlike gFOBTs, FITs are specific for human globin and do not require dietary restriction. Thus, FIT screening is generally associated with higher participation and higher detection rates of both adenomas and CRCs compared to gFOBT screening.[22 23] Furthermore, quantitative FITs offer the opportunity to provide tailored screening by adjusting the positivity cut-off level. This can be used to adjust screening to available resources and colonoscopy capacity.[5 24] A low cut-off increases the detection of advanced neoplasia, but lowers the positive predictive value and specificity thus demanding more colonoscopy resources.[25] No RCT has reported the impact of FIT screening on CRC incidence and mortality. A recent ecological study compared regions in Italy with and without population FIT screening. CRC-specific mortality was 22% lower in areas with a FIT screening program compared to areas without a screening program.[26] The higher uptake and sensitivity of FIT supports the assumption that biennial FIT screening at a low cut-off will have a larger impact than gFOBT on CRC incidence and mortality. Modelling studies suggest that the impact can approach that of colonoscopy if the adherence to multiple rounds is high.[27]

Invasive imaging techniques

Four RCTs showed that a single round of flexible sigmoidoscopy (FS) screening is associated with a reduction in CRC incidence of 18-23% and CRC mortality of 22-31%.[28-31] Similar RCTs evaluating colonoscopy are underway, namely the NordICC (NCT00883792), COLONPREV (NCT00906997), SCREESCO (NCT02078804) and the CONFIRM study (NCT01239082).[32 33] CRC incidence and mortality results from these RCTS are expected between 2025 and 2034.

Colonoscopy is generally considered the gold standard for the detection of colorectal neoplasia. In prospective cohort studies, colonoscopy has been associated with long-term (20 to 30 years) reduction in CRC mortality.[34 35] As

such, some screening programs utilize colonoscopy as the primary screening tool. Other programs prefer a two-step approach, using colonoscopy only for diagnostic clarification in those with a positive first-line less invasive screening test. The latter approach has, for some countries, the advantages of higher screening uptake and lesser demand on limited colonoscopy resources.

Precursor lesions and cancer can be visualized by Computed Tomographic Colonography (CTC), also called virtual colonoscopy. In an average risk population, the per-patient sensitivity of CTC for advanced neoplasia ≥10 mm was 88%.[36] However, this sensitivity decreases for the detection of polyps <10 mm.[37 38] Compared to stool tests, imaging tests such as CTC are more invasive (making them more burdensome) and costlier.

Colonoscopy is considered the primary diagnostic method to evaluate a positive less invasive screening test, whether that test is based on evaluating stool, serum (blood), or colorectal imaging. The demand and capacity for colonoscopy must be taken into account when a country chooses a screening modality.

New screening modalities

Other screening methods have become available for CRC screening. Newer non-invasive tests include DNA, RNA and protein biomarker stool and blood tests. Detection of circulating methylated SEPT9 DNA in blood yielded a CRC sensitivity of 48%, which is at the lower end of the gFOBT range (37–79%).[39 40] Sensitivity for the detection of advanced adenomas was very low (11%).[40] Biomarker stool tests are based on the principle that colorectal neoplasms shed surface cells in the stool. DNA from these cells can be isolated and tested for the presence of mutations and epigenetic changes acquired during carcinogenesis. Stool DNA testing has improved over the last decade. A recent study incorporating FIT with DNA markers, reported a 92.3% sensitivity for CRC and 42.4% for advanced adenomas, which was significantly higher than FIT at a cut-off of 20 µg Hb/g feces (100 ng/ml) (sensitivity for cancer 72%, for advanced adenoma 23%).[41] One cautionary note about this study is the difference in positivity rates between the DNA test and FIT that were used. Both non-invasive tests are meant to select subjects at a higher risk of neoplasia so they can undergo colonoscopy. In other words, a non-invasive test aims to enrich the population undergoing colonoscopy and relieves those at low risk of neoplasia of the burden and risk of colonoscopy. In this particular study, the investigators used a relatively high FIT cut-off and a more "liberal" DNA cut-off. As a result, the number of persons referred for colonoscopy was more than twice as high after stool DNA testing than after FIT. In further comparisons between non-invasive tests, there is a need to set the cut-off of each test at such a level that both tests yield a similar positivity rate, since

this determines colonoscopy demand and thus largely influences the burden and costs of a screening program.

Newer imaging tests include colon capsule endoscopy (CCE) and magnetic resonance colonography (MRC). CCE is a procedure that uses an ingestible capsule with a camera at each end to produce images of the mucosa during intestinal transit. The average sensitivity of second generation CCE (CCE-2) devices for significant findings (≥ 6 mm size, or ≥ 3 polyps irrespective of size) is 86 %. When used as a triage test after a positive FIT to determine who should proceed to colonoscopy, CCE has the potential to reduce the number of colonoscopies performed by 71%. In case of an incomplete colonoscopy, the diagnostic yield of CCE has been reported to be superior to that of CTC for polyps ≥6 mm as well as ≥10 mm with colonoscopy as the gold standard.[42] An MRC study performed in 286 asymptomatic individuals reported a sensitivity of 78.4% for adenomas >6mm.[43] The impact of these new screening modalities on screening uptake and CRC incidence and mortality requires further study.

The efficacy of different screening methods in terms of impact on CRC incidence and mortality is not known. Table 1.2.1 outlines the performance of different screening tests in an average risk population. For the given rates of reduction in CRC incidence and mortality results of a single round, as well as multiple rounds were included in this table. Advanced neoplasia is defined as an adenoma ≥10 mm, or ≥25% villous component, or with high-grade dysplasia or CRC. More recently, attention has also been drawn to the relevance of larger or dysplastic serrated polyps as potential CRC precursors. The accuracy of individual screening methods in detecting these lesions is under study.

Table 1.2.1 Test performance per screening test in asymptomatic, average-risk adults

	gFOBT	FIT	FS	CTC	Colonoscopy
Sensitivity (%) for detecting advanced neoplasia	9 to 24 [44-49]	32 to 53 [44 45 48 50]	90 to 92[a] [51]	88 [36] to 97[44]	88 to 98 [52]
Sensitivity (%) for detecting CRC	13 to 50 [45-47]	79 [53]	90 to 92[a] [51]	100[b] [54]	92 to 99 [51]
Reduction in CRC incidence (%) intention-to-screen	No[c] [20 55]	unknown	18 [55]	unknown	69[d] [56]
Reduction in CRC mortality (%) intention-to-screen	14 to 16 [20]	22[e] [26]	28 [55]	unknown	68[d] [56]

[a] Sensitivity is given for the distal colon; [b] No CRCs were missed by CTC in six screening trials; [c] No reduction in incidence was found in 3 of 4 RCTs included in meta-analysis; [d] Meta-analysis of observational studies, more results expected; [e] ecological study

ORGANIZED AND OPPORTUNISTIC SCREENING

An organized screening program involves a systematic process of inviting a target population to participate in screening and ensuring follow-up of those with a positive screen. An organized program should measure and report on the quality of each step in the screening process. The IARC outlines the following elements for organized screening programs:

- An explicit policy with specified age categories, screening method, and screening interval
- A defined target population
- A management team responsible for implementation
- A health-care team for decisions, care and follow-up of patients with positive screening tests
- A quality assurance structure for every step in the process
- A process for monitoring, evaluating, and identifying cancer occurrence in the population[17]

In organized screening, substantial information technology infrastructure is required to support the program including systems for invitations, recalls, reminders, tracking of screening results, ensuring follow-up and tracking of clinical outcomes such as cancer incidence, mortality and stage.[57] For tracking of screening results, a set of universally applicable CRC screening measures and

indicators has been established.[16] A cancer registry is critical and can be linked to all other relevant databases including laboratories and endoscopic centres.[57] In contrast, opportunistic screening is delivered outside of an organized screening program on an ad hoc basis usually through fee-for-service reimbursement of physicians. Since organized screening focuses on quality assurance, it provides greater protection against the possible harms of screening including over- and under-screening, poor quality, inappropriate use of resources, complications arising from screening, and poor follow-up of those with a positive screen.[58]

The approach to screening in the United States is largely opportunistic. The contributions and quality initiatives from many national bodies has been crucial, including the United States Preventive Services Task Force (USPSTF), an independent volunteer panel of national experts in prevention and evidence-based medicine that reviews evidence and makes recommendations to guide the choice of CRC screening tests.[59] In addition, multiple professional associations have emphasized the importance of colonoscopy quality in the context of CRC screening.[12 60]. Equity of access to screening in the USA remains uncertain, however.[61]

Quality assurance

In 2010, the IARC published the European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis.[13] These guidelines outline targets for key performance indicators for CRC screening including participation, follow-up, and cancer detection rates. For example, the guidelines recommend that invitation coverage in the target population should be high (95%) and that programs should aim for participation rates of at least 65%.[62]

Given that at least 10 years are required to plan, pilot and implement a screening program, the full impact of a nationwide screening program on indicators such as CRC mortality rates requires long follow-up.[63] Therefore, an intermediate measure may be used to evaluate program performance, expressed as the number of persons with advanced neoplasia detected per 1000 invited individuals during the screening interval. This measure takes multiple factors into consideration namely participation rate, positivity rate and the positive predictive value (PPV) for the detection of advanced neoplasia. It is thus a balanced assessment of the overall performance of a screening program. Table 1.2.2 outlines the number of people with advanced neoplasia identified per 1000 invited individuals in those programs that have published their results. There is marked variation across screening tests and within a screening test type for all indicators. Wide ranges for the gFOBT/FIT-based results may be due to the use of more sensitive tests or more stringent criteria for defining test positivity.

Table 1.2.2 Performance of a first round screening program per screening test, based on screening indicators

	gFOBT	FIT	FS[a]	CTC	Colonoscopy
Participation rate (%)	16 – 47	17 – 77	30 – 84	18 – 34	16 - 93
Positivity rate[b] (%)	2.4 – 6.8	1.1 - 13	5.3 – 23	8.6-9.0	4.9 – 11
Advanced neoplasia detection rate[c] (%)	29 – 50	16 - 43	20 –100	54-71	100
Detected advanced neoplasia per 1000 invited individuals[d]	2.1 – 6.3	1.1 – 21	23 – 39	8.8 – 21	14 – 73
References	22 23 64 65	22 23 33 65-73	22 69 74-76	77 78	33 69 77-80

[a] Relative detected advanced neoplasia per 1000 invited individuals is only for the area of the colon examined by FS.
[b] Those with a positive screen were recommended colonoscopy (except when colonoscopy was used as the primary screening test), which enabled the determination of the positive predictive value of the primary screen (the proportion of subjects that during colonoscopy were diagnosed with advanced neoplasia). The uptake of the test was multiplied by the positivity rate and positive predictive value to determine the number of true positives identified with advanced neoplasia per 1000 invited.
[c] Proportion of subjects with a positive primary screening test that were found to have advanced colorectal neoplasia on secondary screening by colonoscopy.
[d] Detection rate per screening round

Cost-effectiveness

Cost-effectiveness studies for CRC screening have concluded that screening is cost-effective compared to no screening.[81-83] Micro-simulation models can help to identify the most appropriate screening strategy given the available resources and budget constraints. The efficiency frontier will identify strategies that are the most effective in terms of life-years gained relative to the cost of the screening strategy.

Cost-effectiveness studies have shown that screening can also be cost-effective in countries with limited financial resources.[84] However, access to and improvement in CRC treatment may be a higher priority than screening in these settings. Using resources to implement population-based screening in a region with no or very limited access to treatment would not be a cost-effective measure.[85 86]

CRC SCREENING PROGRAMS WORLDWIDE

Over the past two decades, the range of CRC screening modalities has expanded, and many population-based programs have been implemented. Nevertheless, large geographic variations remain with respect to implementation of CRC screening (Figures 1.2a-c). As expected screening programs have been more frequently implemented in Western countries with higher CRC incidence and more available resources. Table 1.2.3 shows an overview of screening methods used worldwide grouped into the six WHO regions.

European region

Within Europe, the ASRi rates show a fivefold variation, with lowest rates for men and women in the Balkan countries of Bosnia Herzegovina (30 and 19 per 100 000 respectively) and Albania (13 and 11 per 100 000 respectively). Highest incidence rates in men are found in Slovakia, Hungary and Czech Republic, while highest incidence rates in women are found in Norway, Denmark and The Netherlands.[2] Although CRC mortality rates follow a similar geographic pattern to incidence rates, CRC mortality is also high in some countries with relatively low incidence rates (Moldova, Russia, Montenegro, Poland and Lithuania).[87] A low CRC incidence accompanied by a high CRC mortality can imply limited access to healthcare, and/or suboptimal CRC treatment. It is estimated that in 2015 around 490 000 Europeans will be diagnosed with CRC and 240 000 will die from the disease.[2]

There are large variations among national CRC screening practices in Europe especially since European guidelines for CRC first appeared in 2010.[13] Various screening programs (pilot, opportunistic, or organized) were already in place at that time. There are also considerable differences with respect to financial resources available for screening. The same pertains to colonoscopy capacity, with a more than threefold variation in endoscopy resources across European countries. Taken together, these factors have led to widespread variation. For details by country, see Table 1.2.3.

Most countries in Europe have implemented an organized screening program. Nine countries have an opportunistic program in place, and 16 countries either have or are beginning to implement organized screening (Table 1.2.3). In 2015, 24 out of 28 European Union countries had established or were preparing a nationwide organized or opportunistic CRC screening program.

For instance *Finland, France, Slovenia* and *the United Kingdom (UK)* have completed rollout of their organized programs. In *Belgium, The Netherlands,*

Denmark, Ireland, Italy, Malta, Poland and *Spain*, rollout is ongoing. *Norway, Portugal* and *Sweden* are in the pilot phase.

Some countries have yet to implement a screening program. For instance, the *Greek* Hellenic Society of Gastroenterology released guidelines for CRC screening in 2013. Despite previous low uptake and restricted resources, it recommended colonoscopy as the method of choice for CRC screening. Implementation of an organized program is in a planning phase.

Slovakia has the highest CRC rates in Europe (world standardized ASRi is 43 per 100 000 and ASR mortality 18).[2] However, the country does not have a CRC screening program, despite the Ministry of Health publishing a list of departments performing screening colonoscopy.88 Further work is needed to move forward with a national screening program. No population-based CRC screening program is in place in *Bulgaria, Albania, Bosnia and Herzegovina, Kosovo, Macedonia, Montenegro, Romania, Serbia or Russia*.[89] [90]

Most European countries with an organized program screen by means of a non-invasive stool test, in which previously implemented gFOBT based programs are switching to FIT, such as the *UK* since 2014 and *France* since 2015. Flexible sigmoidoscopy is gradually being introduced in *England*. As of March 2015, about two thirds of screening centers were beginning to offer this one-time-only test to 55 year olds. Countries with an opportunistic program are sometimes faced with low screening uptake. Such is the case in *Austria* and the *Czech Republic,* which has achieved coverage of approximately 25% of the target population.[91] [92] Organized programs may also face uptake issues, such as in *France* and *Croatia*. In *France* participation rates were initially 34.3% in the first two years, with nearly three million people being screened.[93] Active participation by general practitioners increased the participation up to 50% in some parts of the country.[94] The participation rate in *Croatia* after three years of invitations of about 1 million people was only 19.9%.[95]

To match the number of FIT-positives with the available colonoscopy capacity, the FIT cut-off had to be raised in the *Netherlands*, *Scotland* and the *Republic of Ireland*.[96] [97] Compliance to colonoscopy following a positive non-invasive screening test can also be an issue. In *Lithuania* only 52.4% of the FIT-positives undergo colonoscopy; resulting in a very low combined adenoma and CRC detection rate of 1.2% cumulatively over 3 years.[98]

Figure 1.2.1 Overview of screening programs
Regional differences within one country are, except for North-America, not taken into account in these figures. (A) Overview of screening programs in European region. (B) Overview of screening programs in region of the Americas. (C) Overview of screening programs in Western Pacific, South-East Asia and Eastern Mediterranean region.
FIT, fecal immunochemical test for hemoglobin; gFOBT, guaiac fecal occult blood test.

A

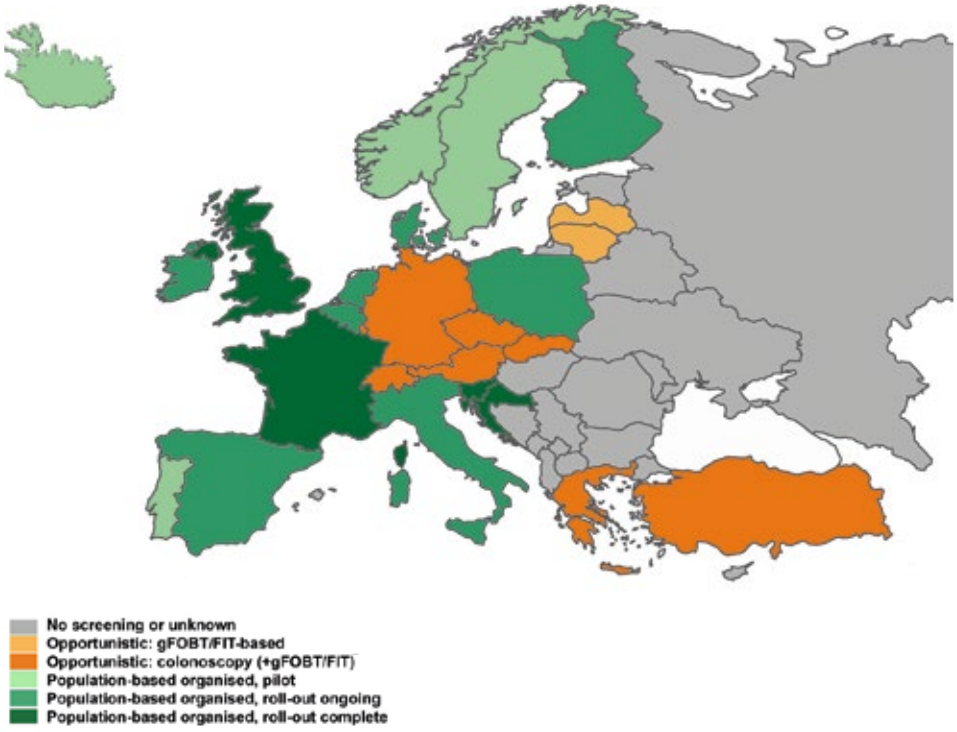
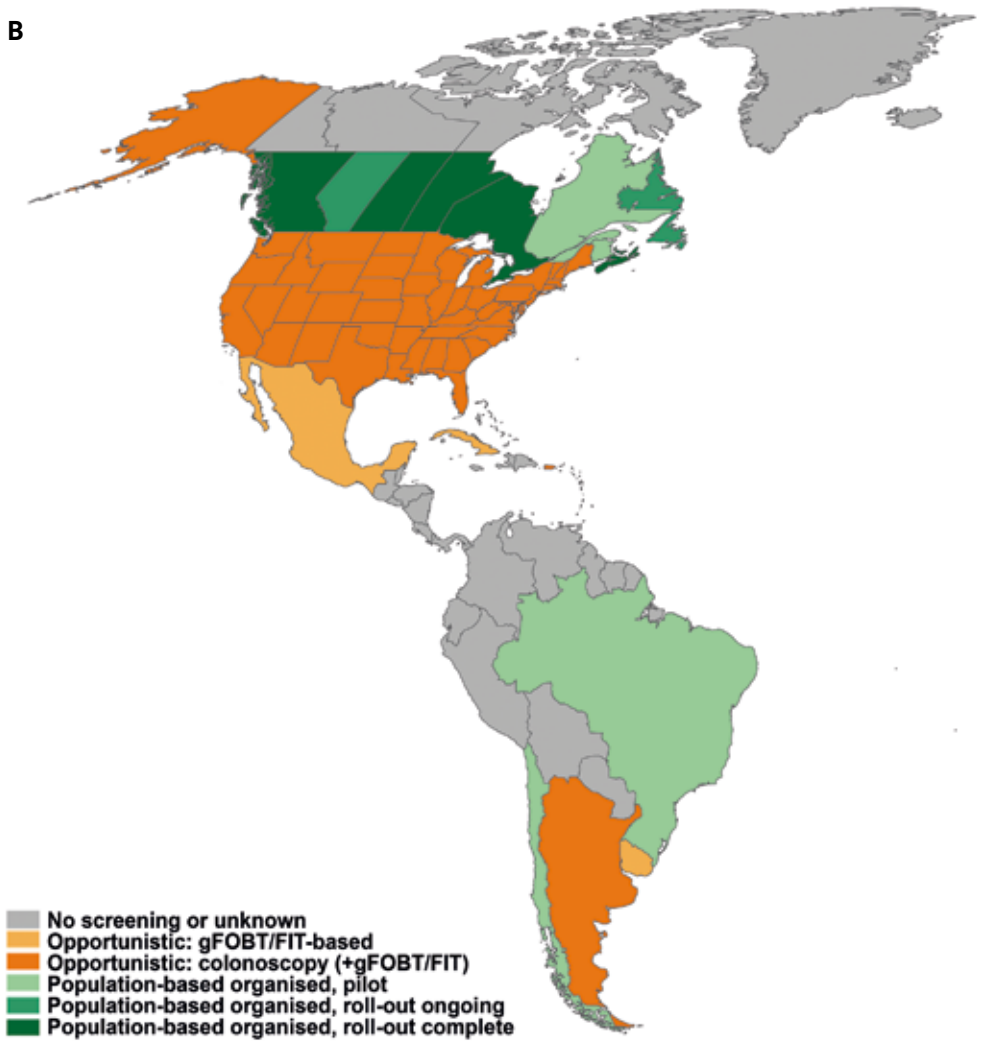


Figure 1.2.1 Continued

B



Region of the Americas

North-America

The ASRi for North America is 26.1 per 100 000 (23 and 30 for women and men respectively).[2] An approximate 136 830 persons have been diagnosed with CRC and 50 310 persons died of the disease in the United States in 2014.[99] An additional estimated 24 400 new cases of CRC were diagnosed in Canada in the same year.[100] Reimbursement for colonoscopy in these Western countries has facilitated the early adoption of opportunistic screening. For example, previous work in Ontario, Canada

demonstrated an increase in colonoscopy use prior to the launch of an organized CRC screening program in the province in 2008.[101]

To date, all 10 *Canadian* provinces have announced, are planning or have implemented organized CRC screening programs.[102] No organized screening programs have been announced in any of the three territories. Special challenges faced by the territories include a lack of resources/facilities and a low population density across vast areas of land. Most provinces are currently using FIT to screen persons aged 50 to 74 years of age at average risk for CRC (Figure 1.2.1B).[102] Ontario's ColonCancerCheck, Canada's first organized CRC screening program, launched province-wide in 2008.[103] Participation in the gFOBT aspect of the program was 29.8% in 2010 to 2011.[103] In Ontario in 2013, 58% of the target pxopulation were up-to-date with CRC screening, taking all screening modalities into account. Early aggregate results from the first round of screening (January 2009 to December 2011) of five other provincial programs (British Columbia, Saskatchewan, Manitoba, Nova Scotia, and Prince Edward Island) showed a much lower participation rate (16%).[104] Opportunistic screening colonoscopy is available to a variable extent in most Canadian provinces.

Colorectal cancer screening in the *United States* is recommended by the USPSTF for persons at average risk (50 to 75 years old) with annual gFOBT, periodic FS, or colonoscopy.[59] Data from the annual Behavioral Risk Factor Surveillance System (BRFSS) survey revealed that approximately 65% of U.S. adults were up-to-date with CRC screening in 2012, with colonoscopy being the most widely used test.[105] The American Cancer Society (ACS) recently reported that CRC incidence and mortality rates significantly decreased over the past decade.[106] This is in particular attributed to CRC screening.[107] Enablers for the success of opportunistic screening seen in the U.S. include the quality initiative supported by the national gastrointestinal societies including the USPSTF, ACS, American Gastroenterological Association (AGA), American Society for Gastrointestinal Endoscopy (ASGE), and National Colorectal Cancer Roundtable (NCCRT) among others. This has played a significant role in the high uptake of screening in the country. On the other hand, organized screening programs in various regions of the country have also been established. Two prime examples include the Kaiser Permanente Northern California (KPNC) Program and the Veterans Health Administration (VHA) program. KPNC participation rates have doubled since 2004 and were 69% in 2010 [108] while screening rates in the VHA among veterans aged 52 or older were already as high as 68% as early as 2001.[109]

Mexico has also launched a CRC awareness campaign focusing on gFOBT screening but without much uptake.

Central and South America

Although numerous Central and South American countries have national guidelines in place for CRC screening, very few national screening programs have been implemented.[110] [111] The current infrastructure in many countries is lacking to support a full organized screening program. In addition, very little is done to raise awareness in most regions including *Venezuela*, *Bolivia*, *Peru*, *Columbia*, and *Costa Rica*. As a result, many of those diagnosed with CRC in these regions are identified after the disease has metastasized.[110]

For example, in *Brazil,* screening colonoscopies for those aged 50 and older have been approved by the Ministry of Health, but uptake remains low.[110] Numerous CRC pilot programs have begun in various municipalities. One example is the program by the Brazilian Association for Colorectal Cancer Prevention, developed in Sao Paulo.[112] Starting in 2006, those 40 years and older at average risk were screened with FIT. Of the 4 567 kits that were distributed between August 2006 and March 2007, 79.7% were returned and analyzed. Positivity rates were approximately 10.7%.[113]

Some success was seen with *Argentina's* CRC screening program in urban areas of the country.[114]The program will next be piloted in rural regions of the country.[110] In *Uruguay*, a CRC pilot program was launched in 1996 for those at average risk. Persons aged 50 or older were screened with FIT and followed up with colonoscopy.73 Between June 1997 and July 2004, 90.1% of the 11 734 persons enrolled in the CRC screening program completed a FIT. Of these, 11.1% had a positive test.

Caribbean

Organized screening for CRC is not routinely performed in the Caribbean. Opportunistic screening with colonoscopy is available on some islands.[111] However, even with effective screening, many Caribbean countries lack the medical facilities to provide appropriate cancer treatment. As such, organized CRC screening is not a priority. For example, in *Jamaica*, screening accounts for 11% of the performed colonoscopies.[115] Similarly, low uptake of CRC screening has been reported in other countries such as *Puerto Rico* and *Cuba*.[116] [117]

African region

A mathematical modelling study showed that in the sub-Saharan African region, screening for CRC by colonoscopy at age 50 in combination with treatment can be considered cost effective.[84] However, the need for population-based CRC screening in the low income countries of Africa is questioned given the overall relatively low burden of disease, the substantial burden of communicable

diseases, and the limited resources.[86] In addition, means by which to identify the target population, availability of colonoscopy and an appropriate number of well-trained specialists are lacking in most regions.[118]

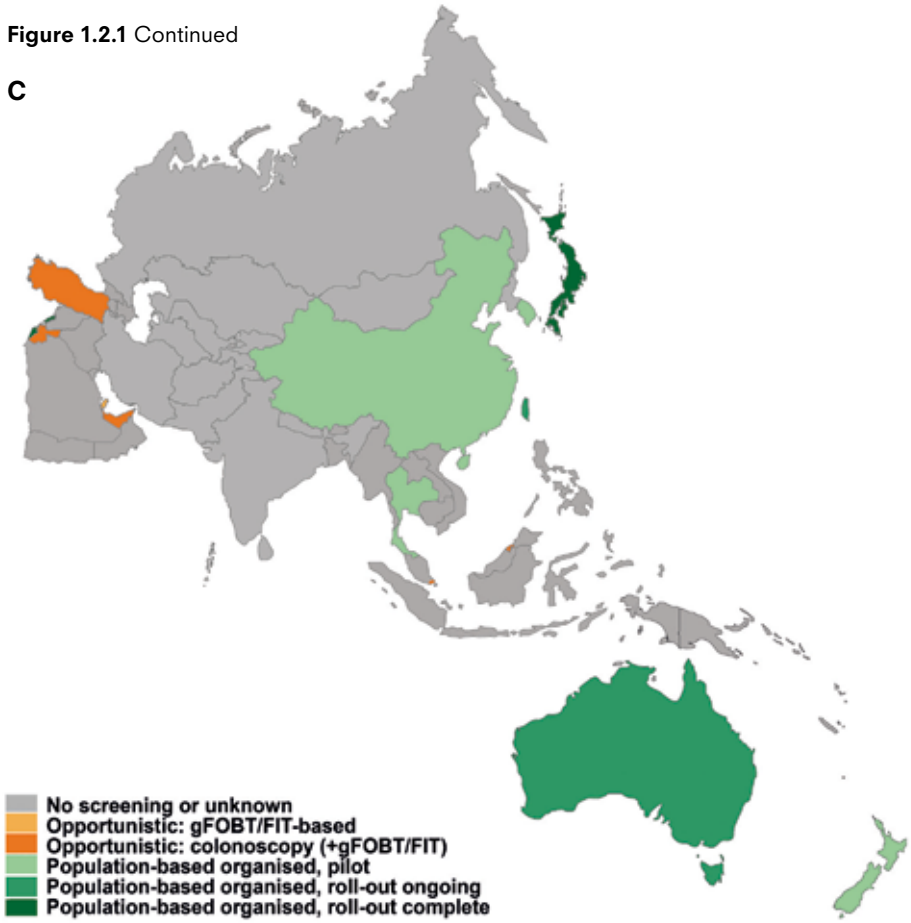
Eastern Mediterranean region

Predictions specific to the Eastern Mediterranean region indicate that generally countries in this part of the world will experience an increase in all cancer mortality of approximately 181 % over the next 15 years.[119] The current CRC incidence is highest in Israel (36 per 100 000), Jordan (26), Kazakhstan (23), Armenia (19), Syrian Arab Republic (16), Lebanon (16) and the State of Palestine (15). CRC mortality rates are highest in Jordan (16 per 100 000), Kazakhstan (13), Armenia (11) and Israel (11). [2]

In *Israel*, an organized program for individuals aged 50-74 has been put in place by the country's four health care providers with the government establishing national oversight of provider activities and quality. Insured persons are approached by their general practitioner (GP) for annual FIT screening. In *Jordan*, the national health authorities have not yet adopted a specific strategy or guidelines for CRC screening despite the high prevalence of the disease in the country.[120] In Middle-Eastern countries that have adopted a Western lifestyle, some opportunistic programs are in place. Physicians in the *United Arabic Emirates* have called for country-wide screening but Abu Dhabi is the only emirate to screen for CRC. The Health Authority Abu Dhabi (HAAD) advises colonoscopy for people over the age of 40.[121] An opportunistic FIT-based pilot study in *Qatar* showed a low colonoscopy follow-up rate among those with a positive FIT of 56%.[122]

Figure 1.2.1 Continued

C



Western Pacific and South-East Asia region

In the Asia Pacific region, CRC incidence varies among regions. The country with the highest CRC incidence in the world is Korea (ASRi 45 per 100 000). Other countries with high CRC incidence in Asia are Singapore (ASRi 34 per 100 000), and Japan (ASRi 32 per 100 000).[2] There is an alarming rising trend in CRC incidence and mortality in Asia, especially in Japan, Korea and China.[14] This trend has been explained by changes in diet and a Westernized lifestyle.[123] In addition, the overall prevalence of advanced colorectal neoplasia in asymptomatic Asian populations is similar to Western populations.[124] The CRC incidence and mortality rates in India (6 and 5 per 100 000 respectively) remain low compared with rising rates in East Asia.[2 125]

Recommendations for CRC screening in the Asia Pacific region have been published[126] and recently updated.[14] The Asia Pacific Colorectal Cancer Working Group recommends CRC screening in regions where the incidence is high, defined

as greater than 30 per 100 000.[14] Recommendations include screening for those 50 to 75 years at average risk with a quantitative FIT as the preferred screening method.[14] Those with a positive test result should be referred for colonoscopy. The guidelines recommend that a clinical risk index can be employed in regions with limited healthcare resources to prioritize screening in those at increased risk. Several studies investigated barriers to CRC screening in different cultural and socio-political contexts in the Asia Pacific region. These barriers included poor knowledge of CRC screening and test characteristics, lack of financial support and lack of health insurance.[127-129]

Several countries in the Asia Pacific region have already developed population-based screening programs. This includes *China*, *Japan*, *Taiwan, Korea* and *Singapore*. In *China,* those aged 40 to 74 are screened with gFOBT and followed up by a digital rectal exam (DRE) and colonoscopy. However, the program is not available to the entire population and the national registry used to track clinical outcomes is estimated to capture only 13% of the country's population, making planning for health care services difficult.[130] Studies have shown that screening uptake in China is low and varies widely.[130 131] In *Japan,* a CRC screening program has been in place since 1992 for national health insurance beneficiaries.[68] Individuals aged 40 to 69 years with national health insurance are offered screening with FIT. In 2010, participation rates for those aged 40 to 69 years were 28.1% for men and 23.9% for women.[132] Nationwide CRC screening was introduced in the *Republic of Korea* in 2004. The National Cancer Screening Program sends invitation letters to the target population at the beginning of the year advising them to get screened with annual FIT, while those with a positive test are offered follow-up with colonoscopy or double contrast barium enema (DCBE).[64 70] Published results from the program show an increase in participation rates since launch (10.5% in 2004 to 21.1% in 2008) and a decline in positivity rates from 8.0% to 6.8%.[64] Opportunistic screening is also available in the region.

In other regions, such as *Hong Kong,* community-based screening programs have been piloted to evaluate the feasibility of a large population-based organized screening program. Similarly in *New Zealand,* a four-year CRC screening pilot began in late 2011 to determine whether a screening program should be rolled out nationally. The pilot results will inform the decision regarding implementation of a national population-based screening program. In *Thailand,* an organized pilot program based on FIT screening has been implemented in April 2011. It focuses on persons aged 50 to 65 years in the Lampang Province.[66] Preliminary results from the pilot show participation rates of 62.9% among the 127 301 persons in the target population.[66]

In *Australia,* a pilot started in 2002. In 2006, the National Bowel Cancer Screening Program (NBCSP) began providing biennial FIT (New Hem Tube) to people turning 55 and 65 years. The program will continue to expand between 2015 and 2020 to fully implement biennial screening for all Australians aged 50 to 74 years.[133] In other regions, such as in *Malaysia,* no organized population-based screening program exists despite published guidelines for CRC screening.[134]

Table 1.2.3 Characteristics of screening programs worldwide, presented alphabetically by country within regions defined by the World Health Organization (last updated May 2015)

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
European region										
Austria	26	9,9	all	opportunistic		gFOBT		1980	40+	12
			all	opportunistic		OC		2005	50+	84-120
Belgium	36,7	11,8								
			Flanders	organized	rollout ongoing	FIT	15 µg Hb/g	2013	56-74	24
			Wallonia and Brussels	organized	rollout complete	gFOBT		2009	50-74	24
Bulgaria	66,6	16								
Croatia	32,9	18,7	all	organized	rollout complete	gFOBT	1/12 slides	2007	50-74	24
Czech Republic	39,9	15,4	all	opportunistic		FIT		2000	50-54	12
			all	opportunistic		OC/FIT		2010	55+	120/24
Denmark	40,5	14,5	all	organized	rollout ongoing	FIT	20 µg Hb/g	2014	50-74	24
Estonia	27,2	12,3								
Finland	23,5	8,3	all	organized	rollout ongoing	gFOBT		2009	60-69	24
France	30	10,2	all	organized	rollout complete	FIT	30 µg Hb/g	2009	50-74	24
Germany	30,9	10,4	all	opportunistic		gFOBT	1/6 slides	1971	50-54	12
			all	opportunistic		OC/gFOBT		2002	55+	120/24
Greece	13,5	7,5	all	opportunistic		OC			50-80	
Hungary	42,3	20,8								

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
Iceland	28,4	7,4		Organized	Pilot/Planning phase					
Ireland, republic of	34,9	12,2	all	organized	rollout ongoing	FIT	45 µg Hb/g	2012	55-74	
Italy	33,9	10,8	all Piedmont / Veneto	organized organized	rollout ongoing pilot	FIT FS	20 µg Hb/g	1982	44-75 58-60	24 once
Latvia	23,7	12,9	all	opportunistic		gFOBT	1/9 slides	2005	50+	12
Lithuania	23,4	13,7	regions	opportunistic		FIT		2009	50-75	24
Luxembourg	31,5	11,2	all	opportunistic		gFOBT/OC		2005	50+	
Malta	31,9	12,2	all	organized	rollout ongoing	FIT		2012	60-64	12
The Netherlands	40,2	13,4	all	organized	rollout ongoing	FIT	47 µg Hb/g	2014	55-75	24
Norway	38,9	13	regions	organized	pilot	OC/FS/FIT		2012	50-64	
Poland	27	14,5	all	organized	rollout ongoing	OC		2000	50-66	120
Portugal	31,7	13,6	center region	organized	pilot	gFOBT		2008		
Romania	26,4	13,4								
Slovakia	42,7	18	all	opportunistic		gFOBT/OC				
Slovenia	37	16,2	all	organized	rollout complete	FIT	1/2 samples at 67 µg Hb/g	2009	50-69	24
Spain	33,1	12,3	regions	organized	rollout ongoing	FIT		2000	60-69	24
Sweden	29,2	10,9	regions	organized	pilot	gFOBT/FIT/OC				
Switzerland	29,4	10,9		opportunistic		gFOBT/OC			50+	

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
Turkey	16,6	10,0	all	opportunistic		FIT/OC		2009	50-74	24/120
United Kingdom	30,2	10,7								
			England	organized	rollout complete	gFOBT	5/6 slides*	2006	60-74	24
					roll-out ongoing	FS		2013	55+	once
					pilot	FIT	20 µg Hb/g	2014	60-74	24
			Scotland	organized	rollout complete	gFOBT	5/6 slides**	2007	50-74	24
					pilot	FIT	80 µg Hb/g	2010	50-74	24
			Wales	organized	rollout complete	gFOBT	5/6 slides**	2008	60-74	24
			Northern Ireland	organized	rollout complete	gFOBT	5/6 slides**	2010	60-74	24
Ukraine	23,4	13,7		unknown						
Region of the Americas										
Argentina	23,8	13	urban areas	organized	pilot	FIT/OC			50-74	12
Bahamas	20,3	10,8		opportunistic		gFOBT/FIT/OC				
Barbados	28,4	14,1		opportunistic		gFOBT/FIT/OC				
Brazil	15,8	8	regions /Sao Paulo	organized	pilot	FIT				
Canada	35,2	10,8								
			Ontario	organized	rollout complete (switching to FIT)	gFOBT	1/6 slides	2008	50-74	24

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
			British Columbia	organized	rollout complete	FIT	1/2 samples at 20 µg Hb/g	2009	50-74	24
			Alberta	organized	rollout ongoing	FIT	15 ug Hb/g	2007	50-74	12 or 24
			Saskatchewan	organized	rollout complete	FIT	20 µg Hb/g	2009	50-74	24
			Manitoba	organized	rollout complete	gFOBT	1/6 slides	2007	50-74	24
			Quebec	organized	pilot/planning phase	FIT		2014	50-74	24
			New Brunswick	organized	pilot/planning phase	FIT		2014		24
			Nova Scotia	organized	rollout complete	FIT	1/2 samples at 300 µg Hb/g	2009	50-74	24
			Prince Edward Island	organized	rollout complete	FIT		2009	50-74	24
			Newfoundland & Labrador	organized	rollout ongoing	FIT		2012	50-74	24
			Yukon	no screening						
			Northwest territories	no screening						
			Nunavut	no screening						
Chile	15	8,6	7 cities	organized	pilot	FIT	20 µg Hb/g			
Cuba	19,7	11,6		opportunistic		FIT				
Jamaica	14,4	7,9		opportunistic		OC				
Martinique	23,9	9,4		organized	rollout complete	FIT	30 µg Hb/g	2007	50-74	24

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
Mexico	7,8	4,1		opportunistic		gFOBT/FIT			50+	
Puerto Rico	24,6	9,6		opportunistic		gFOBT/FS/OC			50-75	
Trinidad /Tobago	23,5	13,1		opportunistic		gFOBT/FIT/OC				
Uruguay	29,5	15,7		opportunistic		FIT	20 µg Hb/g	1997	50+	24
Unites States of America	25	9,2								
			Kaiser Permanente North Carolina	organized		FIT/OC			50-75	12
			Veterans Health Administration	organized		gFOBT/FS/OC			51-75	
Western Pacific, South-East Asia and Eastern Mediterranean region										
Armenia	19,3	11,1		unknown						
Australia	38	9	all	organized	rollout ongoing	FIT			50-74	60
Brunei	25	12		opportunistic		OC				
China	14,2	7,4								
			Hong Kong	organized	pilot	gFOBT/OC		2003	50+	

Country	ASRi	ASRm	Region(s)	Program type	Status of organized program	Type of test	Definition positive test	Starting year	Age range (years)	Screening interval (months)
			Several including Shanghai and Hangzhou regions	organized		gFOBT/DRE + OC		2008	40-74	
Israel	35,9	11,1	all	organized	rollout complete	FIT		1990	50-74	12
Japan	32,2	11,9	all	organized	rollout complete	FIT		1992	40-69	12
Jordan	25,6	15,5		opportunistic		gFOBT/FIT/OC			50+	
Kazakhstan	22,8	12,8		unknown						
Korea, North	21,8	10,7		unknown						
Korea, South	45	12	all	organized		FIT		2004	50+	12
Malaysia	18,3	9,4		no organized screening						
New Zealand	37	15	Waitemata	organized	pilot	FIT	15 µg Hb/g	2011	50-74	
Singapore	33,7	11,8	all	organized		FIT			50+	12
Taiwan			all	organized	rollout ongoing	FIT		2004	50-74	
Thailand	12,4	7,3	Lampang Province	organized	pilot	FIT	200ng/mL	2011	50-65	

* weak positives (1-4 of 6 slides positive) are retested; ** weak positives (1-4 of 6 slides positive) are retested with FIT

CONCLUSION

Colorectal cancer incidence and mortality rates vary widely among continents and within continents. High-quality incidence and mortality data allow understanding of disease and are thus the first essential step for effective cancer control planning. In considering whether to move forward with a CRC screening program, the local impact of the disease relative to other health problems and the capacity to treat the disease adequately should be taken into account. Non-communicable diseases as CRC are rapidly becoming the leading health-care problem in middle-income and low-income countries. This in particular pertains to those countries that are transitioning to Western lifestyles and have aging populations. Therefore, the need to consider implementing CRC screening beyond the countries in which it is currently taking place is likely to increase over time. Most countries with a high CRC incidence however, already have some form of screening in place.

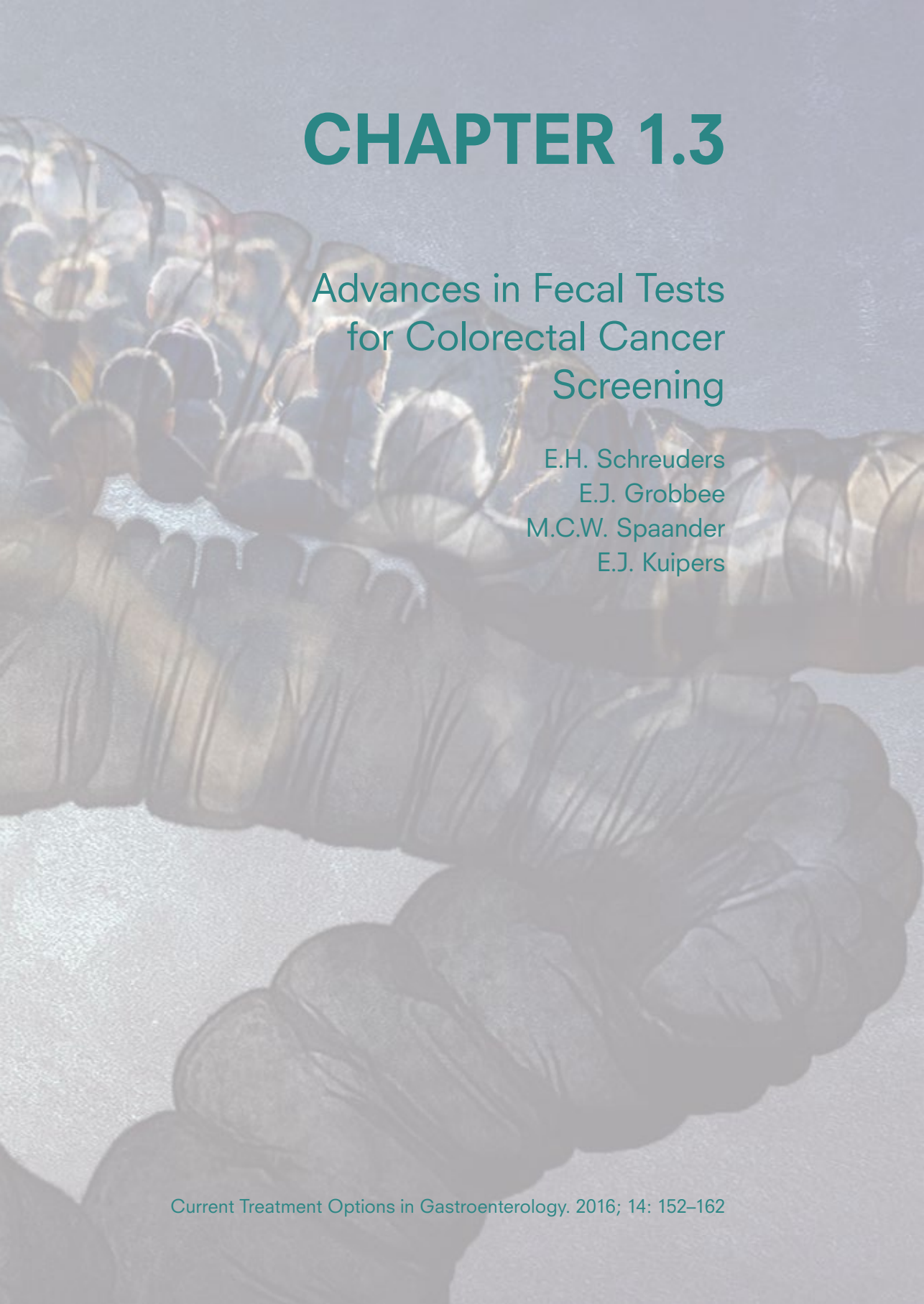
Despite major changes over the past 15 years, there remain many countries without population-based CRC screening despite high CRC incidence and mortality. This is in most cases explained by limitations in resources including colonoscopy capacity, and the organization of structure of health care delivery. Some countries without an existing program already have CRC screening on the agenda. This will likely result in implementation of CRC screening in the coming years.

Most organized CRC screening programs use non-invasive stool tests (FIT or gFOBT), whereas most opportunistic programs are based on endoscopy, in particular colonoscopy. For both screening strategies, levels of screening uptake vary considerably throughout the world. A screening strategy should be chosen carefully to meet the needs of the applicable screening scenario. A comprehensive understanding of the full range of screening modalities and strategies available for CRC screening is needed for appropriate selection of strategies relative to available financial resources and colonoscopy capacity.

The lack of CRC screening in many countries and the low screening uptake in various others provide room for improvement. In countries with a CRC screening program with low uptake levels, targeted actions need to be considered to improve uptake. This may include adaptations to the invitation and follow-up protocol, in particular implementing an active call-recall system. Other measures may include a change to or addition of another screening modality. Professional gastroenterology associations may actively promote such changes in close conjunction with health authorities and screening organizations.

Finally, quality assurance and evaluation is of paramount importance to ensure optimal impact, minimal burden, and balanced use of resources. Therefore, screening measures and quality indicators should be reported, allowing national evaluation and international comparison to improve CRC screening quality.

In conclusion, the global challenge is to evaluate the need for CRC screening in a given jurisdiction or country, and, if indicated, to develop a tailored CRC screening program for which the uptake is high. This is especially necessary for low resource countries that face an increase in CRC incidence, as populations adopt a more Westernized lifestyle.



CHAPTER 1.3

Advances in Fecal Tests for Colorectal Cancer Screening

E.H. Schreuders
E.J. Grobbee
M.C.W. Spaander
E.J. Kuipers

Current Treatment Options in Gastroenterology. 2016; 14: 152–162

ABSTRACT

Colorectal cancer (CRC) forms an important public health problem, especially in developed countries. CRC screening tests can be used to identify asymptomatic individuals with CRC pre-cursors and (early) cancer. Removal of these lesions reduces CRC incidence and prevents CRC related mortality. There are a range of screening tests available, each with advantages and disadvantages. Stool screening tests can broadly be divided into fecal occult blood tests (FOBTs) and molecular biomarker test, such as DNA/RNA marker tests, protein markers and fecal microbiome marker tests. Guaiac fecal occult blood tests (gFOBT) have been demonstrated in large randomized screening trials to reduce CRC mortality. Fecal immunochemical tests (FIT) have superior adherence, usability and accuracy as compared to gFOBT. Advantage of the use of quantitative FITs in CRC screening programs is the cut-off level that can be adjusted. Molecular biomarker DNA tests have shown to detect significantly more cancers than FIT. By combining biomarker DNA tests with FIT, sensitivity for advanced adenomas can be increased significantly. However, it has lower specificity thus demands more colonoscopy resources, is more cumbersome and costly. The adherence has not been assessed in population screening trials. For these reasons, FIT is therefore at present regarded as the preferred method of non-invasive CRC screening. This chapter will review the current status of fecal test based CRC screening.

INTRODUCTION

Colorectal cancer (CRC) forms an important public health problem, especially in developed countries.[1] It ranks third among the most commonly diagnosed cancers worldwide, affecting approximately 1.23 million patients each year.[4] In developed countries, it is the second cause of cancer-related death in men and the third cause in women.[135 136] The high incidence and associated mortality, and the natural history of CRC with slow progression from a premalignant polyp to cancer, makes CRC very suitable for population screening.[6 7] The National Polyp Study within the United States showed that adenoma removal reduced the incidence of CRC by 76–90%. After a median follow-up of nearly 16 years, colonoscopy with removal of adenomas resulted in a 53% reduction in CRC mortality (mortality ratio 0.47, 95% CI 0.26-0.80) compared to the expected CRC mortality rate in the general population.[8 34] Further studies showed screening and prevention of CRC is cost-effective and dependent on strategy also cost-saving.[5]

Various CRC screening tests are available, which can basically be divided into non-invasive stool or blood tests and more invasive imaging or endoscopy procedures. There is no single worldwide-agreed optimal CRC screening method. This results in different approaches in various countries.[137] The choice which screening method should be used is mainly dependent on financial and endoscopy resources, and secondly on the willingness of the population to undergo the primary screening test. As a result of limited resources and population preferences for non-invasive screening, many organized screening programs use a two-step approach. This includes primary screening with a non-invasive fecal test, followed by bowel inspection by means of colonoscopy for individuals that tested positive. For a screening test several test characteristics are necessary. Since screening involves asymptomatic and mostly healthy people, a test should be safe, meaning that the test and screening program should cause no harm. In this light a high test-specificity is preferred, reducing the risks of harm from both unnecessary (follow-up) testing and overdiagnosis. This is contrary to a diagnostic test in a clinical setting, where the high pretest-probability of the disease has often already been established and the disease needs to be confirmed or ruled out. Furthermore, a screening test should be acceptable to the screenee. Adherence rates of those invited for screening are a direct reflection of the acceptance of the test. A screening program with low adherence will by definition not impact on CRC incidence and mortality, irrespective of the characteristics of the test that is offered. Also, practical use and costs of the test need to be taken into account. Fecal tests may differ in positivity rate, and thus number of people referred for colonoscopy. They consequently have different demands for colonoscopy resources. Recently, the variability of fecal tests as CRC screening tool is rapidly increasing and more countries have been implementing CRC screening programs. The large variability and expanding range of fecal tests may impair knowledge of the available screening options.

Therefore, this review aims to give an overview of the recent advances in fecal tests and its use in colorectal cancer screening programs.

FECAL OCCULT BLOOD TESTS

Fecal occult blood tests (FOBTs) detect hemoglobin (Hb) in feces. A range of FOBTs is available; they can be divided into two types: guaiac FOBT (gFOBT) and fecal immunochemical tests (FIT).

Guaiac fecal occult blood tests

Guaiac tests were already available a century ago. They were then used to detect gastric blood loss from peptic ulcer and gastric cancer, conditions that affected large numbers of patients. Guaiac FOBTs were in the 1970s the first widely used FOBT for population-based CRC screening. The gFOBT detects blood by the use of guaiac-impregnated paper to which hydro-peroxidase is added. In contact with heme, the hydro-peroxidase oxygenizes guaiac leading to a blue discoloration. The test-result, i.e. the blue discoloration, is qualitative (positive / negative). The standard gFOBT consists of three paper cards each with two panels, requiring sampling from three separate stools. Guaiac FOBTs can be analyzed with and without hydration. The former has the advantage of a higher sensitivity, however, it also leads to more false-positives.[138] The impact of gFOBT screening on CRC incidence and mortality has been prospectively assessed in several, large randomized trials. These trials demonstrated that repeated annual or biennial gFOBT screening reduces CRC-related mortality by approximately 32-33% and 6-18% respectively.[139-142] The Minnesota trial, which used rehydrated gFOBT, also demonstrated a reduction in CRC incidence.[141] A subsequent meta-analysis reported a pooled 15% reduction in CRC-related death among the three biennial screening trials with gFOBT compared to controls.[20] The Minnesota trial recently after 30 years follow-up reported an overall 27% reduction in CRC mortality.[143]

A main disadvantage of gFOBT is that it does not specifically target human heme. Hydro-peroxidase also reacts with non-human heme present in red meat. This may cause a false-positive test result. Several fresh fruits and vegetables contain peroxidase activity, which may also lead to false-positive test results. Vitamin C may on the contrary block the peroxidase reaction, resulting in false-negative test results.[144] As a result of the dietary restrictions and the need for three different samples on consecutive days, adherence rates of gFOBT-screening are generally poor.[57 144] Furthermore, although gFOBT has a high specificity, its sensitivity is

limited since it does not detect hemoglobin concentrations below approximately 600 µg/g feces.[138] Consequently, adenomatous polyps, precursors of most CRCs, are less likely to be detected as they generally bleed less. The focus on early cancers provides a short window of opportunity, which explains the need for short screening intervals. For these reasons high sensitivity gFOBTs have been designed, with an enhancer to allow detection of lower Hb concentrations.[145] However, these gFOBTs come with a lower specificity making these test less suitable for population-based screening. Due to the low sensitivity and adherence rates, gFOBT screening is associated with a significant proportion of interval cancers.[57] In the Scottish population program, the proportion of interval cancers increased from 31.2% to 58.9% after the first, respectively third screening round.[146] This increase can partly be explained by a decrease in screen-detected cancers over the screening rounds.

Fecal immunochemical tests for hemoglobin (FIT)

Fecal immunochemical tests detect human globin by means of an antibody-based assay. FITs either provide a qualitative result or quantitative result in terms of fecal Hb concentration per gram feces. The latter has the advantage that the selection of cut-off level in population based screening can be tailored to financial and endoscopy resources. There are many different FIT brands available on the market. Figure 1.3.1 shows some of the different FITs. These tests sample different amounts of fecal material, use different amounts of buffers, analytical procedures, and reporting units. They generally present results as Hb concentration in nanograms per ml test buffer. As a result of these differences, the quantitative results of different tests cannot be compared one-to-one. It has therefore been proposed to standardize the reporting units of fecal Hb to µg Hb / gram feces.[19] However, even when using these standardized Hb concentrations, different brand of FITs perform differently in mass screening.[147] These differences apply for both qualitative and quantitative results.[144] Currently, there is no evidence for one FIT to be superior over another.[57]

Figure 1.3.1 FIT brands with different sampling probes, collection tubes, and volume of preservative buffer

At present there are no results from large prospective randomized trials concerning the impact of repeated FIT screening on CRC incidence and mortality. Even so, the current European guidelines recommend FIT screening as the preferred method of fecal occult blood testing.[148] Screening by means of FIT has advantages over gFOBT screening (Table 1.3.1). Firstly, FIT testing requires only one stool sample instead of sampling from three bowel movements. Furthermore, the sampling probe connected to the inside of the lid of the test facilitates test handling (Figure 1.3.1). Together this results in significantly higher adherence rates with higher detection rates of CRC and advanced adenomas.[25] Also, FIT is more sensitive in detecting hemoglobin than gFOBT with reported sensitivities for advanced neoplasia detection of two to three times higher compared to gFOBT.[45] This higher sensitivity for advanced neoplasia would allow prevention of the development of CRC, and thereby potentially decreasing CRC incidence in addition to detecting CRC in an early stage. Lastly, when comparing FIT to gFOBT regarding cost-effectiveness, FIT screening is more cost-effective at any given cut-off. At the same colonoscopy demand, FIT screening led to lower costs and more life years gained than gFOBT.[149]

FIT performance shows variability among different subgroups. Some studies reported a higher sensitivity for left-sided adenomas than right-sided lesions.[150] [151] Also, FIT sensitivity has shown to be higher for aspirin users compared with nonusers.[152] At the same cutoff, men have higher FIT positivity rates than women.[153] This is a reflection of the higher prevalence of advanced neoplasia in men, as well as their more frequent distal location.[154]

Currently, most quantitative FITs have been mainly used with a fixed cut-off, thereby limiting the FIT to a qualitative result (i.e. either positive or negative). Rationale for choosing a specific cut-off greatly depends on the aim of screening and available colonoscopy resources.[149] Using a higher cut-off is of particular interest in situations with limited colonoscopy capacity where screening programs aim for maximal diagnostic yield with restricted resources.[96] A high cut-off also comes with a high positive predictive value. A lower cut-off increases sensitivity for detection of subjects with advanced neoplasia but requires larger colonoscopy resources due to a lower positive predictive value. There is still much to gain concerning the quantitative nature of FIT, as the exact fecal Hb concentration could be of great clinical use. There is evidence that fecal Hb concentration is related to the severity of advanced neoplasia.[155] By adding fecal Hb concentration in predictive models, individuals with the highest risk of advanced neoplasia can be identified.[156] This may also allow for gender-specific approaches. Combining individual fecal Hb concentrations with other risk factors for CRC to base colonoscopy indication on, and not solely a qualitative test result, could possibly improve FIT-screening efficiency. Also FIT could be of clinical importance after the initial positive test result, because the fecal Hb concentration is associated with the risk of a second colonoscopy within one year after screening colonoscopy.[157] Future research in FIT screening should therefore explore the possibilities of incorporating individual fecal Hb levels in CRC screening programs.

Table 1.3.1 Differences between gFOBT and FIT screening in average-risk individuals

gFOBT	FIT
repeat sampling from multiple bowel movements	single sampling from one bowel movement
dietary restrictions	no dietary restrictions
qualitative result	quantitative or qualitative result
semi-automated analysis	automated analysis
sensitivity CRC 31-63%*	sensitivity CRC 69-100%**
specificity CRC 92-96 %*	specificity CRC 92-96%**

*48 49

**151 158 159

DNA- AND RNA-BASED MARKER TESTS

DNA- and RNA-based stool tests aim to detect markers of aberrant DNA or RNA from neoplastic cells. They are based on the principle that colorectal neoplasms shed surface cells in stool. DNA or RNA from these cells can be isolated and tested for the presence of mutations and epigenetic changes acquired during carcinogenesis. DNA- and RNA-based testing is relatively new compared to FOBTs. DNA analysis techniques are developing rapidly and are very sensitive.[160]

DNA markers

A recent study combined FIT with several DNA markers, consisting of molecular assays for aberrantly methylated BMP3 and NDRG4 promoter regions, mutant KRAS, and β-actin (a reference gene for human DNA quantity).[41] This multi-target stool DNA plus FIT test had a significantly higher sensitivity for advanced adenomas (42%) and a somewhat higher sensitivity for CRC (92%) than FIT alone (sensitivity for advanced adenoma 23%, for CRC 72%). However, this increase in detection came against the background of a considerably higher test positivity rate (16% vs 7%). As a consequence, the demand for colonoscopy is more than twice as higher after DNA-FIT testing than after FIT alone, and the DNA-FIT test had a lower specificity for advanced neoplasia compared to FIT, 87% versus 95% respectively. In the light of population based screening and limited colonoscopy resources, the higher positivity rate and lower specificity are important pullbacks. Since colonoscopy resources are in many regions the limiting factor in population screening, it has been advocated to compare non-invasive tests not at a fixed test cut-off, but over a range of positivity rates, allowing a direct comparison between tests at the same positivity rate.[57] Also, the multi-target DNA test requires a full stool sample to be sent in a container, which comes with additional costs and impracticality to an already expensive testing procedure. Furthermore, the multi-target DNA test is recommended once every three years, whereas FIT is offered annually in the Unites States. Cumulative sensitivity, specificity and costs after three years of annual FIT screening would therefore be the fairest comparison before drawing conclusions on superiority of either of the two tests. Lastly, adherence to the multi-target stool DNA test has not yet been investigated. Since adherence is crucial in screening efficacy, this should be evaluated before proceeding to implementation of the test in a screening program.

The performance of DNA tests may differ per CRC subtype, on the grounds that CRC is a heterogeneous disease that can develop via multiple pathways. To detect all CRC subtypes with a screening test, different tumor markers have to be used.[161] [162] A recent study showed that different subtypes are associated with

marked differences in survival. Subjects with tumor markers reflecting a serrated morphology have the highest disease specific mortality (Hazard ratio 2.20).[163] This was compared to subjects with tumor markers reflecting the traditional adenoma-carcinoma sequence (the most predominant tumor). However, the biologic basis for the observed difference remains an important topic for future research.

RNA markers

A microRNA (miRNA) is a small non-coding RNA molecule (containing approximately 22 nucleotides), which functions in RNA silencing and post-transcriptional regulation of gene expression. MiRNAs are thought to be cell-type and disease specific and may be quantified in stool by quantitative real-time polymerase chain reaction (qPCR).[164] Aberrant expression of a specific miRNA may display the effects of a tumor suppressor or oncogene. A large number of studies of single or panels of miRNAs for the detection of CRC have been published recently.[165] A Japanese group showed that the addition of fecal miRNA-106a to FIT testing improved the sensitivity but decreased the specificity of FIT.[166] Since most published studies were based on selected populations, further research in an asymptomatic population should be conducted.

PROTEIN MARKERS

Protein based stool markers focus on either the detection of cancer specific proteins or detection of proteins released from inflamed and/or bleeding tissue. Fecal tumor M2 pyruvate kinase (M2-PK) has received the most attention as a potential cancer specific protein marker. The test is based on the detection of proteins in stool derived from neoplastic colonocytes. A recently reported meta-analysis of studies comparing M2-PK with colonoscopy reported a pooled CRC sensitivity and specificity of 79% and 80%, respectively.[167]

A non-cancer specific protein marker is calprotectin. Calprotectin is a calcium binding protein in granulocytes, macrophages, and epithelial cells. Elevation of calprotectin occurs during intestinal inflammation, including inflammation caused by inflammatory bowel disease. Elevated fecal calprotectin in CRC is suggested to be due to neutrophil shedding from an ulcerated tumor into the intestinal lumen. A large Norwegian CRC screening trial evaluated calprotectin and reported a lower sensitivity and specificity than FIT.[168]

So far, no single protein stool marker has shown to be of adequate accuracy to be considered for population-based CRC screening. A Chinese study investigated the possibility of combining seven biomarkers in a biochip for the

detection of colorectal cancer.[169] The most optimal result was the combination of two biomarkers (TPO, FGF-23) leading to a sensitivity of 0.8 with a specificity of 0.7 for detection of cancer. The use of protein makers in CRC screening, or combining protein markers with FIT, requires further research.

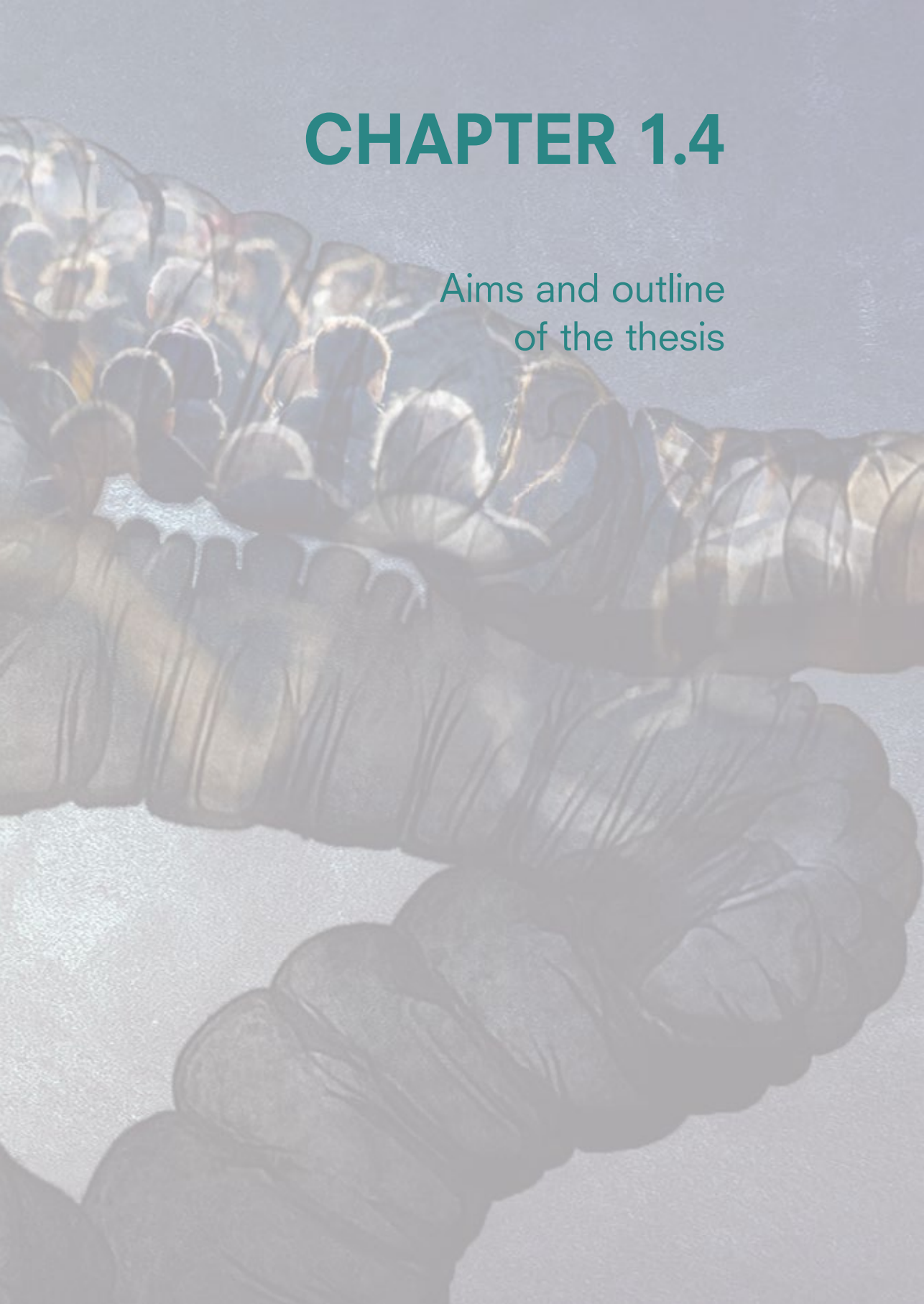
Currently novel molecular tests to analyze stool for a combination of genetic, epigenetic, and protein biomarkers are being developed. The largest is the Molecular Early Detection of Colorectal Cancer (MEDOCC) project. It is a long-term collaborative research between the Netherlands and United States.

HUMAN FECAL MICROBIOME-BASED BIOMARKERS

Recent studies suggest an important role for the gut microbiome in the development of CRC. Patients with CRC have a different gut microbiome than healthy subjects.[170] One of the first bacteria more commonly found in patients with CRC was *Streptococcus bovis*.[171] At present, other bacteria have been identified that play a role in gastro-intestinal cancers, such as *Helicobacter pylori*, and *Fusobacterium nucleatum*.[170] The latter has been of particular interest in colorectal neoplasia, with several studies indicating that *F. nucleatum* in feces is associated with the occurrence of colorectal adenomas and cancer. However, its precise role in this process is poorly understood.[170] It has been suggested that *F. nucleatum* could be useful in detecting serrated polyps.[172] This is relevant as FOBT does not seem sensitive to serrated lesions.[173] The fecal microbiome in CRC screening has largely been unstudied, but measuring the fecal microbiome to identify those at risk of CRC seems promising as a novel screening method. A major advantage of this method of screening could be that non-bleeding lesions are also detected. One study combined the gut microbiome with other risk factors and found that by adding the microbiome the pretest to posttest probability increased, resulting in better identification of subjects with advanced neoplasia.[174] The use of the gut microbiome as a CRC screening tool has great potential. Hence, studies identifying specific microbiota that are associated with CRC are much awaited.

CONCLUSION

There is a wide range of fecal tests for colorectal cancer screening available. The guaiac FOBT was one of the first fecal tests used in colorectal cancer screening. Large trials have shown a significant reduction in CRC-related mortality after screening with gFOBT. However, its use has several limitations when compared to FIT. These limitations include ease of use, adherence, and sensitivity. Also, FIT has the advantage that fecal Hb concentrations can be measured yielding a quantitative test result. Nonetheless, at present FIT is mostly analyzed using a pre-determined cut-off. Aside from FOBTs, to date, only DNA-based stool tests have undergone the full spectrum of development and testing for clinical practice. The multi-target stool DNA test was approved by the FDA in the United States for CRC screening in 2014. So far, biomarker tests such as the multi-target stool DNA test are more expensive than the FOBTs, and come with a relatively low specificity. Furthermore, adherence rates have not been evaluated. Therefore, a sensitive single biomarker of panel of biomarker (stool) test at affordable cost is much awaited. Also, further identification of the gut microbiome could open up new possibilities in CRC screening strategies. Expansion of molecular biomarker screening tests may become imaginable in the future. At the current stage, screening by means of FIT seems the way to go. Requirements in test sensitivity, specificity and costs in order for new molecular biomarker technologies to be cost-effective compared to the FIT should be investigated. Future focus should also be on using FIT quantitatively and incorporating FIT results in risk prediction models to maximize screening benefit and efficacy.



CHAPTER 1.4

Aims and outline of the thesis

This thesis provides strategies and insights that may contribute to optimizing colorectal cancer (CRC) screening. **PART I** starts with an overview of the burden of CRC worldwide and global initiatives to reduce this burden by screening (**Chapter 1.2**). The available screening methods are discussed, which can mainly be divided into non-invasive fecal tests and more invasive colon imaging tests. **Chapter 1.3** concentrates on the recent advances in fecal tests for CRC screening.

In **PART II**, we further focus on fecal occult blood test-based CRC screening. The main two fecal occult blood tests are the guaiac-based fecal occult blood test (gFOBT) and the more recently developed fecal immunochemical test (FIT). Screening by means of gFOBT has been proven to reduce CRC-related mortality, although for FIT these data have not yet become available. We compare the diagnostic accuracy of FIT and gFOBT in average-risk individuals by performing a Cochrane systematic review and meta-analysis (**Chapter 2**). This chapter provides an overview of the test performance characteristics for both types of fecal occult blood tests. The general aim of the subsequent chapters is to explore various aspects of FIT screening and screening strategies. This aim is pursued using important data derived from four rounds of a large prospective population-based CRC screening trial called the 'CORERO'-trial. This study was implemented in 2006 and invited 13,205 asymptomatic average-risk individuals for quantitative FIT screening. FITs can be either qualitative, providing a positive or negative test result, or quantitative, quantifying fecal hemoglobin (Hb) concentrations in feces. Quantitative FIT offers the opportunity of selecting a specific cut-off concentration used to identify candidates for further evaluation by colonoscopy. We aimed to study the effect of various fecal Hb cut-off concentrations and screening age on diagnostic yield of advanced neoplasia, missed lesions, and colonoscopy demand (**Chapter 3**). Because adenomas can bleed intermittently, neoplasia can be missed with the use of one sample of one bowel movement. With the use of two FITs per round, the detection rate of advanced neoplasia may therefore be increased. In **Chapter 4**, we explore the best screening strategy to detect advanced neoplasia in terms of number of FIT samples offered to screenees. In screening, quantitative FITs are invariably used in a dichotomous manner using pre-specified cut-offs. We aim to determine if fecal Hb concentrations of participants with a FIT result below the cut-off could be used to predict future colorectal advanced neoplasia risk (**Chapter 5**).

The aim of **PART III** of this thesis is to provide more insight in factors that are associated with quality in CRC screening, surveillance and colonoscopy. We first focus on online information on CRC screening and surveillance. The internet is

increasingly used for health information and assessing the availability and quality of online information intended for screenees is therefore essential. We evaluated the accuracy, quality, and readability of online information on CRC screening and surveillance in **Chapter 5**. Colonoscopy is the most commonly performed gastrointestinal endoscopic procedure and considered the 'gold standard' investigation of the colon. Aside from primary colonoscopy screening, all CRC screening tests are generally followed by colonoscopy in case of a positive test result. Measuring and optimizing quality of colonoscopy contribute to a higher preventive effect of CRC screening and surveillance. The quality of colonoscopy can be measured by comparing performance parameters i.e. quality indicators of an individual endoscopist or group of endoscopists with predetermined targets. Reduction in variation of quality has also emerged as an important priority for colonoscopy practice. We therefore assessed whether plenary feedback on quality indicators can stimulate improvement of colonoscopy and thereby reduce inter-hospital differences (**Chapter 7**). In **Chapter 8**, we expand our knowledge on quality assurance of surveillance colonoscopies after the removal of adenomas.

In **PART IV**, the main findings of this thesis are summarized and discussed. In addition, the implications for CRC screening and directions for further research are highlighted (**Chapter 9**).

REFERENCES

1. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. Nat Rev Dis Primers 2015;**1**(In Press):15065.
2. GLOBOCAN. Estimated cancer incidence, mortality and prevalence worldwide in 2012 http://globocan.iarc.fr/Default.aspx, 2012.
3. Forman DB, F. Brewster, D.H. Gombe Mbalawa, C. Kohler, B. Piñeros, M. Steliarova-Foucher, E. Swaminathan, R. Ferlay, J. *Cancer Incidence in Five Continents Vol. X*. Lyon: International Agency for Research on Cancer, 2014.
4. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;**127**(12):2893-917.
5. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, et al. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. J Natl Cancer Inst 2009;**101**(20):1412-22.
6. Kuntz KM, Lansdorp-Vogelaar I, Rutter CM, et al. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. Med Decis Making 2011;**31**(4):530-9.
7. Brenner H, Hoffmeister M, Stegmaier C, et al. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 2007;**56**(11):1585-9.
8. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;**329**(27):1977-81.
9. Stryker SJ, Wolff BG, Culp CE, et al. Natural history of untreated colonic polyps. Gastroenterology 1987;**93**(5):1009-13.
10. 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.
11. The Canadian Task Force on Preventive Health Care. Colorectal cancer screening: Recommendation statement from the Canadian Task Force on Preventive Health Care. Can Med Assoc J 2001;**165**(2):206-08.
12. Levin B, Lieberman DA, McFarland B, 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.
13. von Karsa L, Patnick J, Segnan N. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Executive summary. Endoscopy 2012;**44 Suppl 3**:SE1-8.
14. Sung JJ, Ng SC, Chan FK, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. Gut 2015;**64**(1):121-32.
15. World Endoscopy Organization. WEO Colorectal Cancer Screening Committee: WEO/OMED; 2015 [Available from: http://www.worldendo.org/weo-colorectal-cancer-screening-committee.html.
16. Benson VS, Atkin WS, Green J, et al. Toward standardizing and reporting colorectal cancer screening indicators on an international level: The International Colorectal Cancer Screening Network. Int J Cancer 2012;**130**(12):2961-73.
17. International Agency for Research on Cancer. Handbook of Cancer Prevention: cervix cancer screening. Lyon: IARC Press, 2010.

18. Fraser CG, Allison JE, Young GP, et al. 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.

19. 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**(11):810-4.

20. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol 2008;**103**(6):1541-9.

21. Brenner H, Hoffmeister M, Birkner B, et al. Diagnostic performance of guaiac-based fecal occult blood test in routine screening: state-wide analysis from Bavaria, Germany. Am J Gastroenterol 2014;**109**(3):427-35.

22. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 2010;**59**(1):62-8.

23. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;**135**(1):82-90.

24. Wilschut JA, Hol L, Dekker E, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology 2011;**141**(5):1648-55 e1.

25. 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.

26. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut 2015;**64**(5):784-90.

27. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating Test Strategies for Colorectal Cancer Screening—Age to Begin, Age to Stop, and Timing of Screening Intervals: A Decision Analysis of Colorectal Cancer Screening for the U.S. PreventiveServices Task Force from the Cancer Intervention and Surveillance Modeling Network (CISNET). Rockville (MD): Agency for Healthcare Research and Quality (US) 2009;**March**(08-05124-EF-2.).

28. Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet 2010;**375**(9726):1624-33.

29. Schoen RE, Pinsky PF, Weissfeld JL, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. N Engl J Med 2012;**366**(25):2345-57.

30. Segnan N, Armaroli P, Bonelli L, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. J Natl Cancer Inst 2011;**103**(17):1310-22.

31. Holme O, Loberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. JAMA 2014;**312**(6):606-15.

32. Kaminski MF, Bretthauer M, Zauber AG, et al. The NordICC Study: rationale and design of a randomized trial on colonoscopy screening for colorectal cancer. Endoscopy 2012;**44**(7):695-702.

33. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Engl J Med 2012;**366**(8):697-706.

34. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;**366**(8):687-96.

35. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N Engl J Med 2013;**369**(12):1095-105.

36. de Haan MC, van Gelder RE, Graser A, et al. Diagnostic value of CT-colonography as compared to colonoscopy in an asymptomatic screening population: a meta-analysis. Eur Radiol 2011;**21**(8):1747-63.

37. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003;**349**(23):2191-200.

38. Pickhardt PJ, Hassan C, Halligan S, et al. Colorectal cancer: CT colonography and colonoscopy for detection--systematic review and meta-analysis. Radiology 2011;**259**(2):393-405.

39. Whitlock EP, Lin JS, Liles E, et al. 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.

40. Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut 2014;**63**(2):317-25.

41. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014;**370**(14):1287-97.

42. Spada C, Hassan C, Barbaro B, et al. Colon capsule versus CT colonography in patients with incomplete colonoscopy: a prospective, comparative trial. Gut 2015;**64**(2):272-81.

43. Graser A, Melzer A, Lindner E, et al. Magnetic resonance colonography for the detection of colorectal neoplasia in asymptomatic adults. Gastroenterology 2013;**144**(4):743-50 e2.

44. Graser A, Stieber P, Nagel D, et al. Comparison of CT colonography, colonoscopy, sigmoidoscopy and faecal occult blood tests for the detection of advanced adenoma in an average risk population. Gut 2009;**58**(2):241-8.

45. 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.

46. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;**351**(26):2704-14.

47. Lieberman DA, Weiss DG, Veterans Affairs Cooperative Study G. One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon. N Engl J Med 2001;**345**(8):555-60.

48. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;**105**(9):2017-25.

49. Ahlquist DA, Sargent DJ, Loprinzi CL, et al. Stool DNA and occult blood testing for screen detection of colorectal neoplasia. Ann Intern Med 2008;**149**(7):441-50, W81.

50. Chiu H, Lee Y, Tu C, et al. Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clin Gastroenterol Hepatol 2013;**11**(7):832-38.

51. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating Test Strategies for Colorectal Cancer Screening: A Decision Analysis for the U.S. Preventive Services Task Force. Ann Intern Med 2008;**149**(9):659-69.

52. Lieberman DA. Clinical practice. Screening for colorectal cancer. N Engl J Med 2009;**361**(12):1179-87.

53. Lee JK, Liles EG, Bent S, et al. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med 2014;**160**(3):171.

54. de Haan MC, Pickhardt PJ, Stoker J. CT colonography: accuracy, acceptance, safety and position in organised population screening. Gut 2015;**64**(2):342-50.

55. Holme O, Bretthauer M, Fretheim A, et al. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. Cochrane Database Syst Rev 2013;**9**:CD009259.

56. Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. BMJ 2014;**348**:g2467.

57. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol 2013;**10**(3):130-42.

58. Miles A, Cockburn J, Smith RA, et al. A perspective from countries using organized screening programs. Cancer 2004;**101**(5 Suppl):1201-13.

59. U.S. Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008;**149**(9):627-37.

60. Rex DK, Johnson DA, Anderson JC, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. Am J Gastroenterol 2009;**104**(3):739-50.

61. Gupta S, Sussman DA, Doubeni CA, et al. Challenges and Possible Solutions to Colorectal Cancer Screening for the Underserved. J Natl Cancer Inst 2014.

62. Moss S, Ancelle-Park R, Brenner H. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation of screening outcomes. Endoscopy 2012;**44**(3):25.

63. Karsa Lv, Lignini TA, Patnick J, et al. The dimensions of the CRC problem. Best Practice & Research Clinical Gastroenterology 2010;**24**(4):381-96.

64. Choi KS, Lee HY, Jun JK, et al. Adherence to follow-up after a positive fecal occult blood test in an organized colorectal cancer screening program in Korea, 2004-2008. J Gastroenterol Hepatol 2012;**27**(6):1070-7.

65. Levi Z, Birkenfeld S, Vilkin A, 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.

66. Khuhaprema T, Sangrajrang S, Lalitwongsa S, et al. Organised colorectal cancer screening in Lampang Province, Thailand: preliminary results from a pilot implementation programme. BMJ Open 2014;**4**(1):e003671.

67. Lopez-Kostner F, Kronber U, Zarate AJ, et al. [A screening program for colorectal cancer in Chilean subjects aged fifty years or more]. Rev Med Chil 2012;**140**(3):281-6.

68. Saito H. Colorectal cancer screening using immunochemical faecal occult blood testing in Japan. J Med Screen 2006;**13**(1):S6-7.

69. Segnan N, Senore C, Andreoni B, et al. Comparing attendance and detection rate of colonoscopy with sigmoidoscopy and FIT for colorectal cancer screening. Gastroenterology 2007;**132**(7):2304-12.

70. Shim JI, Kim Y, Han MA, et al. Results of colorectal cancer screening of the national cancer screening program in Korea, 2008. Cancer Res Treat 2010;**42**(4):191-8.

71. Tepes B, Stabuc B, Stefanovic M, et al. Faecal immunochemical test-based colorectal cancer screening programme SVIT in Slovenia: pilot phase. Eur J Cancer Prev 2014;**23**(4):235-9.

72. Yeoh KG, Chew L, Wang SC. Cancer screening in Singapore, with particular reference to breast, cervical and colorectal cancer screening. J Med Screen 2006;**13**(1):S14-9.

73. Fenocchi E, Martinez L, Tolve J, et al. Screening for colorectal cancer in Uruguay with an immunochemical faecal occult blood test. Eur J Cancer Prev 2006;**15**(5):384-90.

74. Senore C, Ederle A, Benazzato L, et al. Offering people a choice for colorectal cancer screening. Gut 2013;**62**(5):735-40.

75. Atkin WS, Cook CF, Cuzick J, et al. Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial. Lancet 2002;**359**(9314):1291-300.

76. Weissfeld JL, Schoen RE, Pinsky PF, et al. Flexible sigmoidoscopy in the PLCO cancer screening trial: results from the baseline screening examination of a randomized trial. J Natl Cancer Inst 2005;**97**(13):989-97.

77. Scott RG, Edwards JT, Fritschi L, et al. Community-based screening by colonoscopy or computed tomographic colonography in asymptomatic average-risk subjects. Am J Gastroenterol 2004;**99**(6):1145-51.

78. Stoop EM, de Haan MC, de Wijkerslooth TR, et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. Lancet Oncol 2012;**13**(1):55-64.

79. Lieberman DA, Weiss DG, Bond JH, et al. Use of Colonoscopy to Screen Asymptomatic Adults for Colorectal Cancer. N Engl J Med 2000;**343**(3):162-68.

80. Schoenfeld P, Cash B, Flood A, et al. Colonoscopic screening of average-risk women for colorectal neoplasia. The New England journal of medicine 2005;**352**(20):2061-8.

81. Vanness DJ, Knudsen AB, Lansdorp-Vogelaar I, et al. Comparative economic evaluation of data from the ACRIN National CT Colonography Trial with three cancer intervention and surveillance modeling network microsimulations. Radiology 2011;**261**(2):487-98.

82. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev 2011;**33**(1):88-100.

83. Pignone M, Saha S, Hoerger T, et al. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2002;**137**(2):96-104.

84. Ginsberg GM, Lauer JA, Zelle S, et al. Cost effectiveness of strategies to combat breast, cervical, and colorectal cancer in sub-Saharan Africa and South East Asia: mathematical modelling study. BMJ 2012;**344**:e614.

85. Ginsberg GM, Lim SS, Lauer JA, et al. Prevention, screening and treatment of colorectal cancer: a global and regional generalized cost effectiveness analysis. Cost Eff Resour Alloc 2010;**8**:2.

86. Rabeneck L. HS, Zauber A.G., Earle C. Colorectal Cancer. In: Dean T. Jamison HG, Sue Horton, Prabhat Jha, Ramanan Laximinarayan, and Rachel Nugent, ed. Disease Control Priorities in Developing Countries. 3rd ed ed. Washington, DC: World Bank, 2014:Forthcoming.

87. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013;**49**(6):1374-403.

88. Ministry of Health Slovak Republic. List of departments performing endoscopic screening for colorectal cancer (screening and primary screening colonoscopy) under the Professional Guidelines of the Ministry of Health no. Z06173 / 2011 - CFC for the execution of colorectal cancer screening, published in the Bulletin of the Ministry of Health, Volume 59, figure 19-31, dated September 1, 2011 2001 [Available from: http://www.health.gov.sk/?skrining-kolorektalneho-karcinomu.

89. Sporea IP, A. No colorectal cancer screening program in Romania! Thus, start with opportunistic screening. Rev Med Chir Soc Med Nat Iasi 2014;**118**(3):598-600.

90. Avksentyeva M. Colorectal cancer in Russia. Eur J Health Econ 2010;**10 Suppl 1**:S91-8.

91. Zavoral M, Suchanek S, Majek O, et al. Colorectal cancer screening: 20 years of development and recent progress. World J Gastroenterol 2014;**20**(14):3825-34.

92. Haidinger G, Waldhoer T, Vutuc C. Self-reported colonoscopy screening in Austria. Eur J Cancer Prev 2008;**17**(4):354-7.

93. Leuraud K, Jezewski-Serra D, Viguier J, et al. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer Epidemiol 2013;**37**(6):959-67.

94. Faivre J, Arveux P, Milan C, et al. Participation in mass screening for colorectal cancer: results of screening and rescreening from the Burgundy study. Eur J Cancer Prev 1991;**1**(1):49-55.

95. Katicic M, Antoljak N, Kujundzic M, et al. Results of National Colorectal Cancer Screening Program in Croatia (2007-2011). World J Gastroenterol 2012;**18**(32):4300-7.

96. 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**(3):198-205.

97. van Veldhuizen HH, M.L. Lansdorp-Vogelaar, I. Adjustment to the implementation of the colorectal cancer screening programme in 2014 and 2015. In: (RIVM) TNIfPHatE, ed. Netherlands: RIVM, 2014.

98. Poskus T, Strupas K, Mikalauskas S, et al. Initial results of the National Colorectal Cancer Screening Program in Lithuania. Eur J Cancer Prev 2015;**24**(2):76-80.

99. American Cancer Society CCAG. Cancer Facts & Figures 2014. Atlanta, 2014.

100. Canadian Cancer Society. Canadian Cancer Statistics 2014. Toronto, ON, 2014.

101. Vinden C SS, Rabeneck L. ICES research atlas: Use of large bowel procedures in Ontario. Institute for Clinical Evaluative Sciences. 2004.

102. Canadian Partnership Against Cancer. Colorectal cancer screening in Canada: Program performance results, January 2009-December 2011. Toronto, 2013.

103. Rabeneck L, Tinmouth JM, Paszat LF, et al. Ontario's ColonCancerCheck: Results from Canada's first province-wide colorectal cancer screening program. Cancer Epidemiology Biomarkers & Prevention 2014.

104. Major D, Bryant H, Delaney M, et al. Colorectal cancer screening in Canada: results from the first round of screening for five provincial programs. Curr Oncol 2013;**20**(5):252-7.

105. Centers for Disease Control and Prevention. Morbidity and mortality weekly report: Colorectal cancer screening test use - United States, 2012, 2013.

106. Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. CA Cancer J Clin 2014;**64**(2):104-17.

107. Edwards BK, Ward E, Kohler BA, 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.

108. Levin TR, Jamieson L, Burley DA, et al. Organized Colorectal Cancer Screening in Integrated Health Care Systems. Epidemiol Rev 2011;**33**(1):101-10.

109. Chao HH, Schwartz AR, Hersh J, et al. Improving colorectal cancer screening and care in the Veterans Affairs Healthcare system. Clin Colorectal Cancer 2009;**8**(1):22-8.

110. Latin American Expert Summit for Metastatic Colorectal Cancer. Improving outcomes in the treatment and management of metastatic colorectal cancer in Latin America 2014 [Available from: http://www.angio.org/wp-content/uploads/2014/02/AF-Latin-America-CRC-White-Paper-June2014.pdf.

111. Pan American Health Organization. Cancer in the Americas: Country Profiles 2013. : WHO; 2013 [Available from: http://www.uicc.org/sites/main/files/private/Cancer-Country-Profiles-2013-ENG.pdf.

112. Coy C. Colorectal cancer prevention in Brazil - where are we? Journal of Coloproctology 2013;**33**(3):111-12.

113. Perez RO, Proscurshim I, São Julião GP, et al. Instalação e resultados preliminares de programa de rastreamento populacional de câncer colorretal em município brasileiro. ABCD Arquivos Brasileiros de Cirurgia Digestiva (São Paulo) 2008;**21**:12-15.

114. Instituto Nacional del Cancer. Programa nacional de prevencion y deteccion temprana del cancer colorrectal. [Available from: http://www.msal.gov.ar/inc/images/stories/downloads/publicaciones/equipo_medico/Cancer_colorrectal/guia_CCR_APS_WEB.pdf.

115. Plummer JM, Mitchell DI, Ferron-Boothe D, et al. Colonoscopy in central Jamaica: results and implications. West Indian Med J 2012;**61**(6):610-4.

116. Gonzalez RS. Cancer Screening: Global Debates and Cuban Experience MEDICC Rev 2014;**16**(3-4):73-77.

117. Lopez-Charneco M, Perez CM, Soto-Salgado M, et al. Correlates of colorectal cancer screening among Hispanics: Results from the 2008 Puerto Rico behavioral risk factor surveillance system survey. P R Health Sci J 2013;**32**(2):68-75.

118. Lambert R, Sauvaget C, Sankaranarayanan R. Mass screening for colorectal cancer is not justified in most developing countries. Int J Cancer 2009;**125**(2):253-56.

119. Dey S, Soliman AS. Cancer in the global health era: opportunities for the Middle East and Asia. Asia Pac J Public Health 2010;**22**(3 Suppl):75S-82S.

120. Omran S, Barakat H, Muliira JK, et al. Knowledge, experiences, and barriers to colorectal cancer screening: a survey of health care providers working in primary care settings. J Cancer Educ 2015;**30**(1):53-61.

121. Health Authority Abu Dhabi. Cancer Screening Recommendations: HAAD; 2015 [Available from: http://www.haad.ae/simplycheck/tabid/131/Default.aspx.

122. John A, Al Kaabi S, Dweik N, et al. Emerging role for colorectal cancer screening in Asian countries. Trop Gastroenterol 2014;**35**(1):21-4.

123. Sung JJ, Lau JY, Goh KL, et al. Increasing incidence of colorectal cancer in Asia: implications for screening. Lancet Oncol 2005;**6**(11):871-6.

124. Byeon JS, Yang SK, Kim TI, et al. Colorectal neoplasm in asymptomatic Asians: a prospective multinational multicenter colonoscopy survey. Gastrointest Endosc 2007;**65**(7):1015-22.

125. Pathy S, Lambert R, Sauvaget C, et al. The incidence and survival rates of colorectal cancer in India remain low compared with rising rates in East Asia. Dis Colon Rectum 2012;**55**(8):900-6.

126. Sung JJ, Lau JY, Young GP, et al. Asia Pacific consensus recommendations for colorectal cancer screening. Gut 2008;**57**(8):1166-76.

127. Cai SR, Zhang SZ, Zhu HH, et al. Barriers to colorectal cancer screening: a case-control study. World J Gastroenterol 2009;**15**(20):2531-6.

128. Sung JJ, Choi SY, Chan FK, et al. Obstacles to colorectal cancer screening in Chinese: a study based on the health belief model. Am J Gastroenterol 2008;**103**(4):974-81.

129. Koo JH, Leong RW, Ching J, et al. Knowledge of, attitudes toward, and barriers to participation of colorectal cancer screening tests in the Asia-Pacific region: a multicenter study. Gastrointest Endosc 2012;**76**(1):126-35.

130. Goss PE, Strasser-Weippl K, Lee-Bychkovsky BL, et al. Challenges to effective cancer control in China, India, and Russia. Lancet Oncol 2014;**15**(5):489-538.

131. Kang LNQ, R.L. Cancer screening and prevention in China. . Cancer Control 2014;**8**:131-33.

132. Center for Cancer Control and Information Services NCC. Cancer statistics in Japan '13. Japan: Foundation for Promotion of Cancer Research, 2013.

133. Australian Government. Health Portfolio Budget Statements. Budget Related Paper No. 1.10. Department outcomes. Outcome 1: Population health. 2014.

134. Malaysian Society of Gastroenterology & Hepatology. Screening for colorectal cancer in Malaysia. Consensus/clinical practice guidelines. 2001 [Available from: http://www.acadmed.org.my/view_file.cfm?fileid=209

135. Jemal A, Bray F, Center MM, et al. Global cancer statistics. CA Cancer J Clin 2011;**61**(2):69-90.

136. Center MM, Jemal A, Smith RA, et al. Worldwide variations in colorectal cancer. CA Cancer J Clin 2009;**59**(6):366-78.

137. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.

138. 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.

139. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;**348**(9040):1467-71.

140. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;**348**(9040):1472-7.

141. 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.

142. Lindholm E, Brevinge H, Haglind E. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. Br J Surg 2008;**95**(8):1029-36.

143. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;**369**(12):1106-14.

144. Young GP, Symonds EL, Allison JE, et al. Advances in Fecal Occult Blood Tests: the FIT revolution. Dig Dis Sci 2015;**60**(3):609-22.

145. Allison JE, Sakoda LC, Levin TR, 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.

146. Steele RJ, McClements P, Watling C, 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.

147. Chiang TH, Chuang SL, Chen SL, 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.

148. Von Karsa LP, J. Segnan, N. (eds). *European guidelines for quality assurance in colorectal cancer screening and diagnosis*. Lyon, 2010.

149. Wilschut JA, Habbema JD, van Leerdam ME, et al. Fecal occult blood testing when colonoscopy capacity is limited. J Natl Cancer Inst 2011;**103**(23):1741-51.

150. Haug U, Kuntz KM, Knudsen AB, et al. Sensitivity of immunochemical faecal occult blood testing for detecting left- vs right-sided colorectal neoplasia. Br J Cancer 2011;**104**(11):1779-85.

151. Chiu HM, Lee YC, Tu CH, et al. Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2013;**11**(7):832-8 e1-2.

152. Brenner H, Tao S, Haug U. Low-dose aspirin use and performance of immunochemical fecal occult blood tests. JAMA 2010;**304**(22):2513-20.

153. Kapidzic A, van der Meulen MP, Hol L, et al. Gender Differences in Fecal Immunochemical Test Performance for Early Detection of Colorectal Neoplasia. Clin Gastroenterol Hepatol 2015;**13**(8):1464-71 e4.

154. Ferlitsch M, Reinhart K, Pramhas S, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA 2011;**306**(12):1352-8.

155. Digby J, Fraser CG, Carey FA, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. J Clin Pathol 2013;**66**(5):415-9.

156. Stegeman I, de Wijkerslooth TR, Stoop EM, et al. Risk factors for false positive and for false negative test results in screening with fecal occult blood testing. Int J Cancer 2013.

157. Grobbee EJ, Kapidzic A, van Vuuren AJ, et al. Second-Look Colonoscopies and the Impact on Capacity in FIT-Based Colorectal Cancer Screening. Am J Gastroenterol 2015;**110**(7):1072-7.

158. Chen YY, Chen TH, Su MY, et al. Accuracy of immunochemical fecal occult blood test for detecting colorectal neoplasms in individuals undergoing health check-ups. Advances in Digestive Medicine 2014;**1**(3):74-79.

159. Hernandez V, Cubiella J, Gonzalez-Mao MC, et al. Fecal immunochemical test accuracy in average-risk colorectal cancer screening. World J Gastroenterol 2014;**20**(4):1038-47.

160. Dickinson BT, Kisiel J, Ahlquist DA, et al. Molecular markers for colorectal cancer screening. Gut 2015;**64**(9):1485-94.

161. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007;**50**(1):113-30.

162. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010;**138**(6):2088-100.

163. Phipps AI, Limburg PJ, Baron JA, et al. Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology 2015;**148**(1):77-87 e2.

164. Ren A, Dong Y, Tsoi H, et al. Detection of miRNA as non-invasive biomarkers of colorectal cancer. Int J Mol Sci 2015;**16**(2):2810-23.

165. Hollis M, Nair K, Vyas A, et al. MicroRNAs potential utility in colon cancer: Early detection, prognosis, and chemosensitivity. World J Gastroenterol 2015;**21**(27):8284-92.

166. Koga Y, Yamazaki N, Yamamoto Y, et al. Fecal miR-106a is a useful marker for colorectal cancer patients with false-negative results in immunochemical fecal occult blood test. Cancer Epidemiol Biomarkers Prev 2013;**22**(10):1844-52.

167. Uppara M, Adaba F, Askari A, et al. A systematic review and meta-analysis of the diagnostic accuracy of pyruvate kinase M2 isoenzymatic assay in diagnosing colorectal cancer. World J Surg Oncol 2015;**13**:48.

168. Hoff G, Grotmol T, Thiis-Evensen E, et al. Testing for faecal calprotectin (PhiCal) in the Norwegian Colorectal Cancer Prevention trial on flexible sigmoidoscopy screening: comparison with an immunochemical test for occult blood (FlexSure OBT). Gut 2004;**53**(9):1329-33.

169. Wang HP, Wang YY, Pan J, et al. Evaluation of specific fecal protein biochips for the diagnosis of colorectal cancer. World J Gastroenterol 2014;**20**(5):1332-9.

170. Narayanan V, Peppelenbosch MP, Konstantinov SR. Human fecal microbiome-based biomarkers for colorectal cancer. Cancer Prev Res (Phila) 2014;**7**(11):1108-11.

171. Kuipers EJ, de Jong A. [Gastrointestinal disorders and Streptococcus bovis bacteremia] Gastro-intestinale aandoeningen en Streptococcus bovis-bacteriemie. Ned Tijdschr Geneeskd 1990;**134**(28):1337-9.

172. Ito M, Kanno S, Nosho K, et al. Association of Fusobacterium nucleatum with clinical and molecular features in colorectal serrated pathway. Int J Cancer 2015;**137**(6):1258-68.

173. van Doorn SC, Stegeman I, Stroobants AK, et al. Fecal immunochemical testing results and characteristics of colonic lesions. Endoscopy 2015;**47**(11):1011-7.

174. Zackular JP, Rogers MA, Ruffin MTt, et al. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res (Phila) 2014;**7**(11):1112-21.

PART TWO

Fecal occult blood test based colorectal cancer screening

Chapter 2	Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals *Submitted to Cochrane Database of Diagnostic Test Accuracy Reviews*
Chapter 3	Effects of increasing screening age and fecal hemoglobin cut-off concentration in a colorectal cancer screening program *Clinical Gastroenterology and Hepatology. 2016 Aug 24 (Epub ahead of print)*
Chapter 4	Comparison of multiple rounds one versus two-sample fecal immunochemical test-based colorectal cancer screening *Submitted*
Chapter 5	Fecal hemoglobin concentrations predict future advanced colorectal neoplasia in long-term population-based FIT-screening *Submitted*

CHAPTER 2

Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals

E.H. Schreuders*
E.J. Grobbee*
A.H.C. van Roon
L. van Dam
A.G. Zauber
I. Lansdorp-Vogelaar
G.J.M. Borsboom
E.W. Steyerberg
M.E. van Leerdam
M.C.W. Spaander
E.J. Kuipers

*Both authors contributed equally to this manuscript.

Submitted to Cochrane Database of Diagnostic Test Accuracy Reviews

ABSTRACT

Background

Worldwide, many countries have adopted colorectal cancer (CRC) screening programmes, often based on faecal occult blood tests (FOBTs). These FOBTs fall into two categories based on technique and detected blood component: qualitative guaiac-based FOBTs (gFOBT) and the more recently developed faecal immunochemical tests (FIT) that can be both qualitative and quantitative. Screening by means of gFOBT has been proven to reduce CRC-related mortality. The effectiveness of FIT screening in decreasing CRC-related mortality has not yet been studied in large long-term prospective randomised controlled trials.

Objectives

The primary objective of this review was to compare diagnostic accuracy of gFOBT and FIT screening for detecting advanced colorectal neoplasia in average-risk individuals.

Search methods

We searched MEDLINE, EMBASE, Cochrane Library, BIOSIS Citation Index, and Science Citation Index Expanded until March 21st 2016. We also searched references lists and PubMed-related articles of included studies to identify additional studies. We did not restrict studies based on language, date, or whether data had been collected prospectively or retrospectively.

Selection criteria

Only studies that provided the number of true positives, false positives, false negatives, and true negatives for gFOBT and/or FIT with colonoscopy as the reference standard were included. Two types of studies were included; those in which all participants underwent both the index test and the reference standard (type I); and those in which only participants with a positive index test underwent the reference standard while the negatives were followed for at least one year for development of interval carcinomas (type II). The target population were asymptomatic, average-risk individuals undergoing colorectal cancer screening were included.

Data collection and analysis

Two authors independently selected studies for inclusion and collected

the data from each study. In case of doubt a third author made the final decision. We used a bivariate and HSROC statistical model to obtain summary estimates of sensitivity and specificity and summary ROC curves.

Main results

A total of twenty-three type I studies involving 85,403 participants were included, reporting on a total of thirty-two faecal occult blood tests. Six studies evaluated gFOBT, thirteen studies evaluated FIT, and four studies included both gFOBT and FIT. Twenty-one studies reported advanced neoplasia as outcome, and nineteen studies reported on colorectal cancer. The cut-off for positivity of FIT varied between 2.4 to 50 µg Hb/g faeces. The summary curve estimated by the HSROC model showed that FIT had a higher discriminative ability than gFOBT for advanced neoplasia ($p=0.002$), and colorectal cancer ($p=0.025$).

We included nineteen type II studies reporting a total of twenty-three tests involving 1,495,344 participants. Overall six gFOBT studies, ten FIT studies, and three studies combining both gFOBT and FIT were included. The cut-off for positivity of FIT varied between 2.4 to 10 µg Hb/g faeces. The summary curve estimated by the HSROC model showed that FIT had a higher discriminative ability than gFOBT for colorectal cancer ($p<0.001$).

Authors' conclusions

FIT is superior to gFOBT in detecting advanced neoplasia and colorectal cancer in average-risk individuals. The specificity of both tests is similar. These results strongly support current guidelines for implementing FIT-based CRC screening programs and the switch from gFOBT to FIT testing for existing programs.

BACKGROUND

Based on the Wilson & Jungner criteria published in 1968 and updated by the World Health Organization in 2008,[1 2] screening is justified when (1) a disease is common and associated with significant morbidity or mortality; (2) screening tests are sufficiently accurate in detecting early stage disease, are acceptable to invitees, and are feasible in general clinical practice; (3) treatment after early detection by screening improves prognosis relative to treatment after usual diagnosis; and (4) the potential benefits outweigh the potential harms and costs of screening. Colorectal cancer (CRC) screening fulfils all of these criteria.

There are various methods for CRC screening. These vary in level of supporting evidence, effectiveness, test-related burden, costs and willingness of target subjects to undergo screening. The screening modalities for CRC broadly fall into two categories; (a) faecal tests (i.e., faecal occult blood tests and faecal DNA testing), and (b) partial or full structural exams (i.e., flexible sigmoidoscopy, colonoscopy and computed tomography-colonography (CTC)). Colonoscopy can be used as the reference standard for those with a positive screening test or as a primary CRC screening tool.

Stool blood tests are conventionally known as faecal occult blood tests (FOBTs), which are used as a two-step testing approach in CRC screening (i.e., positive test result requires further examination with visualization of the colon, predominantly by means of colonoscopy). FOBT screening is based on the principle that a large proportion of colorectal neoplasia bleed microscopically before any clinical signs or symptoms become noticeable. Bleeding tends to be intermittent, and blood is distributed unevenly in the stool. The concept of detecting CRC by testing for blood in the stool is based on the observation that cancers bleed because of disruption of the normal mucosa. The amount of blood increases with the size of the polyp and/or the stage of the cancer.[3-6] In general, the amount of faecal haemoglobin tends to be absent or low in those without neoplasia, higher for those with advanced adenomas, and highest for those with CRC.[7] Faecal occult blood testing detects a higher proportion of CRCs and a lower proportion of advanced adenomas, since CRCs tend to have a more constant bleeding pattern and give rise to higher amounts of blood in stools than advanced adenomas, which are believed to bleed more intermittently. In this way FOBT screening identifies those individuals who are most likely to have advanced neoplasia. Therefore, it should be followed by visualization of the colon and rectum. Colonoscopy is considered the gold standard for detection of advanced neoplasia with high sensitivity and specificity (both above 90%) and has the advantage that (adenomatous) polyps and early CRCs

can be removed during the same procedure. A meta-analysis of the accuracy of colonoscopy (performed for various indications), reported that the pooled miss rate for adenomas ≥10mm was 2%, for ≥5-10mm 13%, and for 1-5mm 26%.[8]

FOBTs fall into two categories based on the detected component of blood: guaiac-based FOBTs (gFOBT) and the more recently developed faecal immunochemical tests (FIT) for haemoglobin.

Guaiac-based faecal occult blood test

Guaiac-based FOBTs enable detection of occult blood in stool through the pseudo-peroxidase activity of haem. However, peroxidase also reacts with non-human haem present in red meat. Also, several fresh fruits and vegetables contain peroxidase activity, which may lead to false-positive test results. Vitamin C may block the peroxidase reaction, resulting in false-negative test results. Guaiac FOBTs may detect bleeding from any site in the gastro-intestinal (GI) tract, including the stomach, as haem remains relatively stable during transport through the GI-tract.[9] The usual gFOBT protocol consists of three test cards, each containing two panels. The screenee is instructed to collect two faecal samples from three consecutive bowel movements yielding a total of six stool panels. Applying a hydrogen peroxide reagent to the faeces on the guaiac material in the panel leads to oxygenation of guaiac, which in turn leads to a blue colour change when haem is present. A panel is considered positive if such coloration appears.[10] The number of positive panels for referral to colonoscopy varies between screening programs. In most programs, a single positive panel is sufficient for referral, however in others the number of positive panels is set at five out of six. In this case, less positive panels imply renewed gFOBT testing. Prior to faecal sampling, individuals are asked to restrict their diet and medication as this might affect the number of false-positive and false-negative test results.

The sensitivity and specificity of gFOBT screening varies widely due to the variation in type of test (brand), instructions for stool collection, number of stool samples per screening round, the use of non-hydrated or rehydrated stool samples, double reading of the test, the number of positive panels used to refer a screened person for colonoscopy, and the interval between successive screening rounds. In some trials, rehydrated gFOBT has been studied; rehydration reduces the false negative rate (improves sensitivity) while increasing the false positive rate (reduces specificity).[11 12]

Guaiac FOBTs are the only stool tests for which there is evidence of efficacy from four prospective, randomised controlled trials (RCTs). These trials from the USA, United Kingdom, Denmark and Sweden demonstrated that multiple rounds of annual or biennial gFOBT screening can reduce CRC-related mortality by

approximately 13-33%.[11 13-15] The American trial, which used rehydrated gFOBT, also demonstrated a reduction in the incidence of CRC.[11] A subsequent meta-analysis reported a pooled 15% reduction in CRC-related mortality among the three biennial screening trials with gFOBT compared to controls.[16] The American trial recently reported an overall reduction in CRC mortality of 27% after 30 years of follow-up.[17] The efficacy of gFOBT screening in reducing CRC-related mortality is limited due to a limited sensitivity for detecting CRC and low sensitivity for detection of advanced adenomas.[18] Furthermore, the process of analysing gFOBTs is time consuming and is faulted by the possibility of inaccurate processing and evaluation.[19]

Faecal immunochemical test

FITs have several technological advantages compared to guaiac based screening. FIT specifically targets human globin, a protein that along with haem constitutes the haemoglobin molecule. Therefore, FITs only detect human blood, in contrast to the gFOBT which can falsely detect other substances. For this reason, FITs are less subject to interference by dietary factors and medication. Studies have suggested that NSAID or aspirin use increased the sensitivity of FIT without a decrease in specificity.[20 21] In addition, FITs are more specific for lower GI-tract bleeding since globin is degraded by digestive enzymes in the upper GI-tract. This improves their specificity for neoplasia in the colon and rectum. The sample collection for most FIT variants is less demanding than for gFOBT-sampling, both in terms of requiring a single sample and less direct handling of stools (smear cards for gFOBTs vs brush/spatula for FIT testing). Furthermore, FIT screening does not require dietary restrictions. Both qualitative and quantitative FITs have been developed and are described below.

Qualitative FITs

Qualitative tests require a manual interpretation of test results as positive or negative. There is a range of such tests on the market. They often use immunochromatographic technology, and allow for simple, office-based analysis. Since qualitative FITs provide dichotomous test results and thresholds for a positive test differ between brands, test performances differ.[22 23] However, like gFOBT, inter-observer variations in interpretation of test results may influence performance.

Quantitative FITs

Quantitative FITs on the contrary can be analysed automatically, quantifying the amount of haemoglobin found in the stool sample. One advantage of quantitative FITs in CRC screening programs is that the cut-off level (i.e. the amount of haemoglobin above which the test is considered positive and individuals are referred for follow-up examination) can be adjusted. This allows the number of FIT-positives to be matched with the available resources for further investigation, in particular colonoscopy capacity.[24] Quantitative FITs have further important

advantages over qualitative FITs due to the use of automated analysis. This automation removes inter-observer variation in interpretation of test results, improves reproducibility, and allows for high-throughput testing. Nevertheless, studies suggest variable performance of different brands of quantitative FITs, even when the standardized same cut-off is used.[22]

To date, there are no long-term prospective randomized data that demonstrate that FIT is superior to gFOBT in terms of reducing CRC-related mortality. However, a recent ecological study compared regions in Italy with and without population FIT screening. CRC-specific mortality was 22% lower in areas with a FIT screening program compared with areas without a screening program.[25] With this review, we aim to compare the diagnostic test accuracy measures of gFOBT and FIT screening in order to answer the question "Can gFOBT be replaced by FIT for primary CRC screening?". In order to answer this question, an overview of the test performance characteristics for both types of FOBT will be provided.

Target condition being diagnosed

FOBT screening primarily aims at early detection of bleeding colorectal neoplasia, since only bleeding lesions can be detected by stool blood tests. CRC screening in general aims at lowering CRC mortality by early detection of CRC and lowering CRC incidence by removal his pre-malignant lesion i.e., adenomatous polyps.

Index test(s)

The tests under evaluation are two FOBTs: gFOBT and FIT. More detailed information about the tests and the methods of execution have been previously described. FIT can be both quantitative as well as qualitative, the latter does not report individual faecal Hb concentrations.

Alternative test(s)

There are several alternative tests that can be used for CRC screening purposes. These tests vary in the level of supporting evidence, attendance, effectiveness, and test-related burden, costs. Alternative screening modalities usually considered as effective CRC screening tools include flexible sigmoidoscopy, colonoscopy, computed tomography-colonography and, more recently, capsule endoscopy, faecal DNA testing and serum molecular markers.[26 27]

Rationale

In the Western world, many countries have adopted a CRC screening program, often based on FOBT.[28] Screening by means of gFOBT has been proven to reduce CRC-related mortality. The results on effectiveness of FIT screening in decreasing CRC-related mortality are not yet available. The main explanation for

this is that many countries have already implemented a CRC screening program. In addition, decisions on the optimal screening test have to be based on data about the sensitivity and specificity, existing RCT results, and modelling.[29]

OBJECTIVES

The primary objective of this review is to compare the diagnostic test accuracy of gFOBT and FIT screening for detecting advanced colorectal neoplasia in average-risk individuals.

Investigation of sources of heterogeneity
We aimed to investigate the following sources of heterogeneity.

A. Heterogeneity related to characteristics of the study population (i.e. sex, age limits, ethnicity, selection of invitees (identified from general practitioner records or population registers), cancer stage, distribution and cancer location.
B. Heterogeneity related to the number of FOBTs performed per screening round
C. Heterogeneity related to the cut-off value used for FIT or the number of positive panels used to refer a gFOBT-screened person for colonoscopy.

Due to reasons described later in this review, analysis for heterogeneity could not be performed for all of these factors.

METHODS

Criteria for considering studies for this review
Types of studies: Two different types of studies were included and categorized in this review:

Type I studies: All (randomised, comparative) accuracy studies in which all participants underwent both the index test and the reference standard. Diagnostic case-control studies were considered inappropriate for this review because such studies are likely to overestimate diagnostic performance.[30] Moreover, literature suggests that measures of accuracy may vary with the prevalence and stage-distribution of the target condition.[31] For instance, the sensitivity of a test will often vary according to the severity of the detected disease (e.g. advanced CRCs are more easily detected with FOBTs than early stage tumours). For these reasons, we did not include case-control studies in this review.

Type II studies: All (randomised, comparative) accuracy studies in which all participants with a positive index test were referred for the reference standard and all participants with a negative index test were followed for at least one year to identify development of interval carcinomas. Only data from the first screening round were included for analysis.

Participants

Asymptomatic average-risk individuals aged 40 years and above were considered as representative for a CRC screening program. Study participants included subjects volunteering for a medical health check-up (including CRC screening), as well as individuals identified from population registers, and general practitioner or managed care organisation records.

Index tests

The index test was either gFOBT or FIT (both qualitative and quantitative) as described previously in the Background section.

Comparator tests

Studies were included regardless of whether they made comparisons with other CRC screening modalities.

Target conditions

The primary target condition was CRC, which was defined as the invasion of malignant cells beyond the lamina muscularis mucosa. Patients with an intra-mucosal carcinoma or carcinoma in situ were classified as having high-grade dysplasia.[32] The secondary target condition was advanced neoplasia, which included CRC and advanced adenomas. An advanced adenoma was defined as an adenoma with a greatest dimension of at least 10 mm, or an adenoma with ≥ 25% villous component, and/or high-grade dysplasia.[32] For each included study, we assessed whether these definitions were applied. If another definition was adopted in a study, we stated this in the characteristics of included studies table.

Reference standards

Studies were included for this review if colonoscopy was used as the primary reference standard. Only in case of an incomplete colonoscopy, CTC or double contrast barium enema (DCBE) was accepted as reference. Furthermore, in type II studies participants with a negative index test result had to be followed for at least one year to assess the development of interval carcinomas. Interval carcinomas were defined as CRC diagnosed in an FOBT negative screenee in the period between two successive FOBT screening rounds.[33] If a study did not use this definition, we stated this in the characteristics of included studies table.

Exclusion criteria

We excluded studies where more than 5% of the population consisted of high-risk individuals. High-risk individuals were defined as patients with a history of CRC; subjects with a personal history of adenoma(s); individuals scheduled for diagnostic colonoscopy because of hereditary CRC syndromes or a positive family history of CRC; symptomatic subjects with complaints suspicious for CRC such as rectal blood loss, changed bowel habits, or weight loss; and all patients with a history of inflammatory bowel disease. We also excluded studies in which a positive gFOBT test result needed to be confirmed by a positive FIT test result or vice versa. We excluded studies in which less than 75% of the participants with a positive FOBT underwent colonoscopy or in case of an incomplete colonoscopy, CTC or DCBE.

Search methods for identification of studies

Electronic searches

To identify appropriate studies, the Trial Search Coordinator of the Cochrane Colorectal Cancer Group in collaboration with the Medical School Library of the Erasmus MC conducted a literature search by using the electronic databases Medline, Embase, the Cochrane Library, BIOSIS, SCI-expanded, Pubmed publisher and Google scholar.[34] The Embase, Medline and Biosis searches were run in OVID. There were no restrictions on date or language of the articles being reviewed. Native speakers related to our departments and personal acquaintances translated articles written in languages other than English. The searches were developed using the Boolean term 'AND' between the topics colorectal cancer and faecal occult blood test. To cover the topic of colorectal cancer, the searches were developed by searching MeSH and/or EmTree terms colorectal neoplasms, colorectal cancer and large intestine tumour, and the text words: colorectal, rectal, rectum, colon*, cancer*, carcinoma*, adenocarcinom*, neoplas*, tumor*, tumour*, polyp* and adenom*. To cover the topic of faecal occult blood testing, the search was developed by searching MeSH and/or EmTree terms occult blood, immunochemistry and feces analysis, and the text words; faecal, fecal, feces, faeces, stool*, occult blood, occult blood test*, FOBT*, gFOBT*, FIT*, immunochem*, immunological*, guaiac*, fecal immunochem*, faecal immunochem* and test*. In exploratory searches, we identified articles that used gFOBT brand names without explicitly mentioning either gFOBT or FOBT in the title, abstract or MeSH terms. We have therefore incorporated most of the gFOBT brand names in our literature search. This brand name issue was not present when searching for articles related to the FIT. Differing from the previous published search strategy in our protocol, we did not include the brand name "colorectal" as this yielded too many irrelevant results. Initial searches were conducted on January 31st, 2015. We performed a

further search on March 21st, 2016. Those results have been incorporated in the review. Five studies were added to 'Studies awaiting classification' and will be incorporated into the review at the next update.

Searching other resources

The references of all included relevant studies were hand-searched for additional trials. In addition, we searched for articles citing the relevant studies included in the review. We defined relevant included articles as any included article that was published within the 5 years preceding our search. Furthermore, we searched PubMed for Related Articles of the most relevant included articles. We examined the first 20 results from 'PubMed Related Articles' after sorting by publication date from newest to oldest. We also contacted principal investigators of the included articles to clarify aspects of methods and results, and ask for any unpublished data in the area of FOBT characteristics, where necessary.

Data collection and analysis

Selection of studies

Two reviewers (EJG and EHS) independently assessed whether the titles and abstracts were eligible for further reading. After this initial retrieval, all selected articles were read entirely. Disagreements about including a study for this review were resolved through discussion with a third reviewer (AvR). All studies that did not meet the inclusion criteria, as ascertained in reading the full article, were listed in a separate table with reasons for exclusion. The reference management software EndNote X7 was used for the selection process.

Data extraction and management

Data were extracted from those trials that fulfilled the inclusion criteria. The data that were extracted for both advanced neoplasia and CRC were:

- Positivity rate (PR), i.e. the proportion of participants having a positive index test result.
- True positives (TP), i.e. participants having a positive index test result, followed by detection of advanced neoplasia by means of the reference standard.
- False positives (FP), i.e. participants having a positive index test result, but no advanced neoplasia when assessed with the reference standard.
- True negatives (TN), i.e. participants having a negative index test result, and no advanced neoplasia during colonoscopy for Type I studies and no interval CRC identified during follow up for Type II studies.
- False negatives (FN), i.e. participants having a negative index test result, and advanced neoplasia during colonoscopy for Type I studies and interval CRC identified during follow up for Type II studies.

The analyses only include the main outcome measures sensitivity and specificity (which were derived from TP, FP, TN, FN). For CRC we included data from both type I and II studies, and for advanced neoplasia we included data from type I studies only. For both study types, the extracted data were merged into separate 2 x 2 tables (containing TP, FP, FN, TN). We excluded non-interpretable test results and FOBT-positives who refused to undergo the reference standard from the 2x2 table and in consequence from the meta-analysis. If data were lacking in a specific article we contacted the principal investigators to ask for the original data and/or tried to reconstruct the aforementioned cell frequencies from the information that was published. If this was not successful, we excluded the study. The numbers of participants analyzed in our review is stated in the characteristics of included studies table. The data presented in the 2 x 2 table were used to conduct meta-analysis on sensitivity and specificity. For type II studies, only data regarding CRCs were generally available during follow-up. Therefore, for type I studies we were able to extract data for both advanced adenomas and CRC, but for type II studies we extracted data for CRC only. We extracted data for all possible cut-offs. Since the concentration used for cut-off [ng Hb/ml buffer in the device] is unique to the device or system and cannot be compared with other devices, cut-offs were transformed to the internationally accepted unit of µg Hb/g faeces.[35] All data were extracted independently by two reviewers (EJG and EHS).

Assessment of methodological quality

Two authors (EJG and EHS) independently assessed the quality of each individual study using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.[36] We excluded some of the questions from the QUADAS-2 tool in case they were not applicable. Details of each study are described in the Characteristics of Included studies table. If data was not specified in the article, this was mentioned. When authors did not respond also manufacturers were contacted to retrieve additional details about the test used if needed.

Statistical analysis and data synthesis

Descriptive analysis

The descriptive analysis provides an overview of all available studies. Tables were split by gFOBT or FIT, and by type I or type II studies. For all study types, the following test characteristics were extracted into 2x2 tables: TP, FP, TN, and FN. The extracted data were entered into Review Manager 5. Study-specific estimates and exact 95% confidence intervals (CI) of sensitivity and specificity were obtained and displayed in forest plots per test type. Different symbols were used per test type, in order to create a clear overview of between-test variability.

Inferential statistics

In secondary analyses, we compared the performance of the gFOBTs and versus the FITs. We complied with the methods and techniques introduced and explained in chapter 10 of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[37] We started with an exploration of the study-specific sensitivities and specificities that were extracted from the included studies using RevMan software. Based on the available reported sensitivities and specificities, we used the bivariate model, and the Rutter and Gatsonis Hierarchical summary receiver-operator curves (HSROC) model[38] to explore differences between tests and identify potential sources of heterogeneity. When at least three studies were found in which both gFOBT and FIT were compared with the reference standard we analysed these as a separate group from studies in which only one of the two tests was compared. When less of these studies were found, we allowed that these contributed an observation to both series of tests. These studies were then included twice in the data set. We also calculated initial summary estimates of sensitivity and specificity when it turned out that more than three studies using common cut-off points for test positivity were available. If enough studies appeared per test brand with the same outcome parameters, we performed meta-analyses.

The HSROC model was used to analyse sensitivities and specificities for type I and type II studies separately, including all studies. For the quantitative FIT where 2x2 tables for multiple cut-offs were available, we used the cut-off as advised by the manufacturers for this analyses.

We investigated the effect of cut-off by carrying out subgroup meta-analyses for cut-offs where sufficient data were available. As some studies reported 2 x 2 data for more than one cut-off, this analysis allowed us to include all of the available data. To analyse the sensitivities and specificities from the various studies and test types that used the same cut-off, we used the bivariate model as introduced by van Houwelingen et al., extended by Reitsma et. al., and explained in the Cochrane Manual Chapter 10. [37 39 40] The bivariate model was fitted using proc nlmixed from SAS version 9.2 conforming to the examples of SAS syntax given in chapter 10 of the Manual. We primarily compared the accuracy of gFOBT and FIT tests, employing a model that had the same between study variance/covariance matrix in these two test types. In order to include studies that had zero counts in any of the four cells we added 0.5 to cells containing no observations. Model output provided confidence and prediction region parameters and summary estimates of test accuracy measures per test type (gFOBT or FIT).

Investigations of heterogeneity

The I2 statistic[41] was not calculated as it doesn't account for heterogeneity explained by phenomena such as positivity threshold effects, it is therefore not routinely used in Cochrane DTA reviews. The magnitude of observed heterogeneity was depicted graphically by the prediction ellipse. We planned to address heterogeneity by adding covariates of interest to the HSROC model. The factors that we aimed to include in our heterogeneity analyses are described in our objectives. However, there are several caveats to keep in mind:

A. Heterogeneity related to gender was assessed as percentage of male participants; investigation of heterogeneities within other population characteristics was not feasible due to lack of information provided in individual studies. Nevertheless, only (studies with) average-risk individuals, as defined in our protocol, were included. The well-defined criteria resulted in homogenous studies according to prediction ellipses. Investigation of heterogeneities within cancer stage, distribution and cancer location was not feasible due to lack of information provided in individual studies.
B. Heterogeneity related to the number of tests and/or the number of stools per screening round was assessed.
C. Summary estimates of sensitivity and specificity for the two most common cut-off levels to refer screenees for further evaluation by the reference standard (i.e. cut-off value for FIT in µg/ml or the number of positive panels for gFOBT) were performed. Heterogeneity related to the quantitative or qualitative nature of FIT was assessed.

Sensitivity analyses

Sensitivity analyses were performed in which the QUADAS-items were used to identify studies that scored differently on certain quality items to determine the effect of poor study quality on the overall results. The impact of each study was tested by removing each one from the analysis separately and recalculating the summary estimates.

Assessment of reporting bias

Investigation of publication bias in diagnostic test accuracy studies has proven to be problematic, because many studies are done without ethical approval or study registration.[42-44] Therefore, identification of studies from registration until final publication of the results is not possible.[42] Furthermore, funnel plot-based tests that are commonly used to detect publication bias in reviews of randomised controlled trials, have been shown to be misleading for diagnostic test accuracy reviews such as ours. Therefore we did not assess reporting bias.

RESULTS

Results of the search

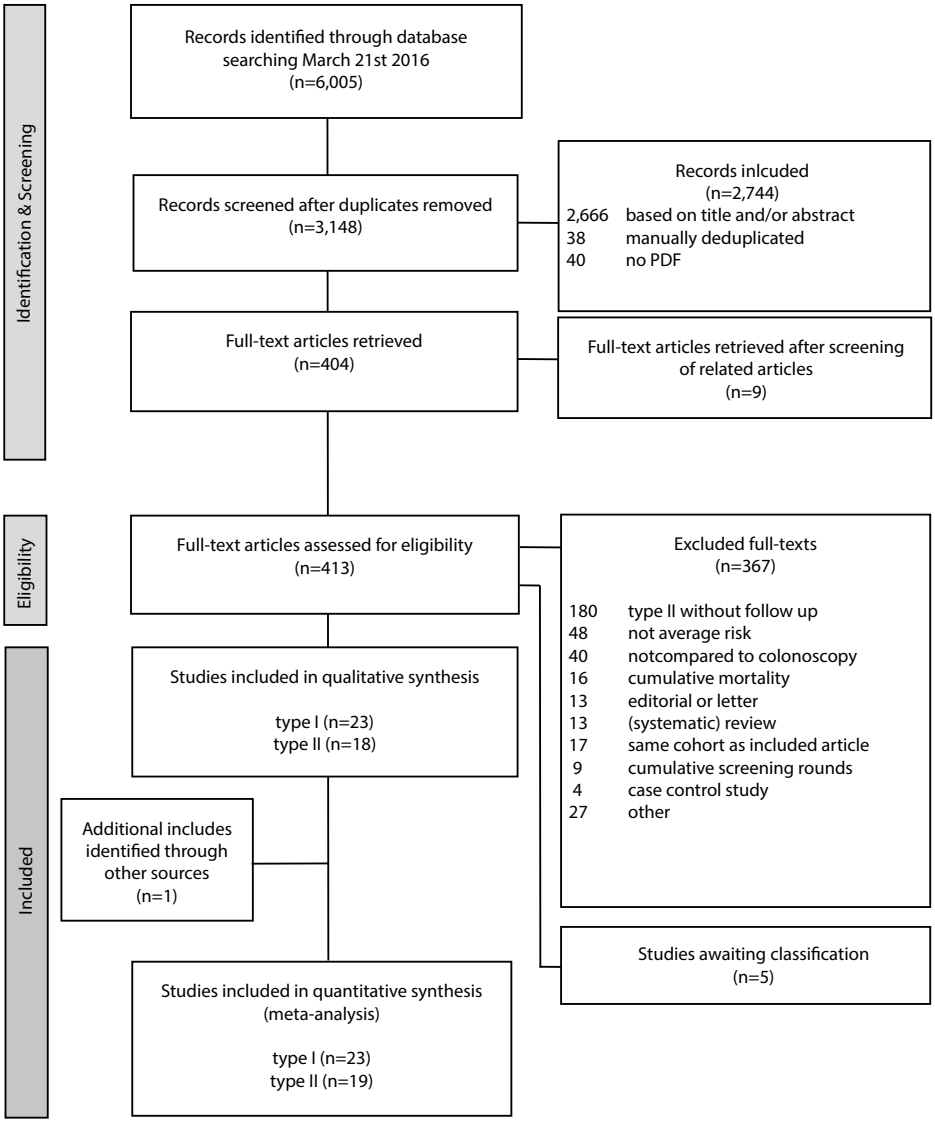
The search identified 6,005 titles, of which 3,148 remained after removal of duplicates. Of these, 2,666 were excluded on the basis of title and abstract. Manual removal of duplicates for the remaining articles resulted in 38 additional duplicate articles. From 40 articles, the PDF could not be retrieved, even after trying to contact the authors. Full articles were retrieved for 404 titles. After hand-searching the references of all included articles and PubMed related articles of included studies, nine additional articles were identified and fully assessed.

An updated search in March 2016 identified 391 additional records. From this updated search, a total of 355 records were excluded based on title and/or abstract. Full-texts were assessed for 36 articles of which 31 have been excluded with reasons and 5 study reports were added to 'Studies awaiting classification' (Figure 2.1).

In total, 413 full-text articles were assessed for eligibility, of which 367 articles were excluded because they met one of the exclusion criteria or were otherwise assessed as ineligible for the following reasons:

- In 180 studies, only FOBT-positive subjects had undergone the reference standard without follow-up of FOBT-negatives;
- Forty-eight studies did not focus on average-risk subjects;
- Forty studies had not used colonoscopy as the first choice of reference standard;
- Sixteen studies only provided data on cumulative mortality over multiple screening rounds making it impossible to determine absolute numbers of advanced neoplasia detected per screening round;
- In 13 studies the full-text was a letter or editorial, and 13 were reviews;
- Seventeen articles encompassed the same cohort as an already included article;
- Nine articles summarized the results of multiple screening rounds and separate data-extraction of the first round was not possible;
- Four articles focussed on digital FOBT where a stool sample was obtained by digital rectal examination;
- Twenty-seven articles were excluded for various other reasons (Figure 2.1).

Figure 2.1 Flowchart of search and included studies.



A total of 41 studies were included. One additional study (Faivre 2004) was included after contacting the author to obtain data on another article to be included.[45] We combined two articles (Brenner 2013, Haug 2011) for gFOBT and FIT results.[46 47] Both studies analysed the same population of 3,077 patients. For the Haug's study the authors provided results for different FIT cut-offs to allow direct comparison with other studies using the same cut-off. The 15 excluded cases in the original article for the analysis about left/right sided lesions were included in our analysis after contacting the principal investigators.

A total of 23 type I studies were included: five were performed in the United States, five in Germany, four in Taiwan, two in the Netherlands, two in Japan, and the remaining studies were performed in France, Spain, South-Korea, China and Hong Kong. Seven studies compared more than one test and of those there were six studies in which participants had undergone more than one index test; resulting in 91,971 test evaluations with a total of 32 separate tests in 85,403 participants. Overall, six gFOBT studies, thirteen FIT studies, and four studies combining both gFOBT and FIT screening were included for this review. The earliest study was published in 2000 (Nakama 2000), with the majority being published between 2008 and 2013. For all but one study (Nakama 2000) advanced neoplasia was the main outcome. Twenty papers separately described the numbers of detected CRC and advanced adenoma. In two studies, no CRC was detected. For eleven studies, data for a cut-off of 10 µg Hb/g faeces could be retrieved, and for eight studies a cut-off of 20 µg Hb/g faeces was used. For all included gFOBT studies a positivity criterion of at least one positive panel was used. All but two studies used a single stool sample for FIT-testing and for all but two gFOBT-studies three consecutive stools were used (Table 2.1).

A total of nineteen type II studies were included: four were performed in France, three in Italy, two in Japan, two in the Netherlands, and the remaining studies were performed in Taiwan, Denmark, Israel, Ireland, Scotland, Finland, Spain (Tenerife), and Germany. Four studies compared more than one index test with each participant undergoing one FOBT. In total they reported on a total of 23 tests in 1,495,344 participants. Overall, six gFOBT, ten FIT, and three combination studies were included. Out of the studies combining gFOBT and FIT, two studies randomized participants and one study performed both tests in all participants. The earliest study was published in 1987 (Kronborg 1987), with the majority being published in 2002 to 2014. All studies had at least one year of follow-up, with a maximum of four years of follow-up. Eleven of nineteen included studies had exact two years of follow-up. For four studies, data with a cut-off of 10 µg Hb/g faeces could be retrieved, and for four studies (five FITs) data with a cut-off of 20 µg Hb/g faeces. All but one study used a single stool sample for FIT-testing and for all gFOBT-tests, three consecutive stools were used (Table 2.2).

Table 1 Overview test characteristics per study for type I studies

Study	Country	Test brand	gFOBT	FIT	FIT method	FIT 10 µg	FIT 20 µg	Other cut-off	AN	CRC	Nr of stools
Alhquist 2008 [48]	USA	Hemocult	+						+	+	3
		Hemocult Sensa	+						+	+	3
Brenner G 2010 [49]	Germany	HemoCARE (gFOBT)	+						+	+	3
		ImmoCARE-C (FIT)		+	Qualitative			unknown	+	+	3
Brenner H 2013 [46]	Germany	Hemoccult	+						+	+	1
Chen 2014 [50]	Taiwan	OC-light		+	Qualitative	+			+	+	1
Cheng 2002 [51]	Taiwan	OC Hemodia		+	Qualitative	+			+	+	1
Chiu 2013 [52]	Taiwan	OC-light		+	Qualitative	+			+	+	1
Cruz-Corrrea 2007 [53]	USA	Hemocult II	+						+		3
de Wijkerslooth 2012 [54]	Netherlands	OC-sensor		+	Quantitative	+	+		+	+	1
Graser 2009 [55]	Germany	Brand not specified	+						+	+	3
		FOB-Gold		+	Quantitative			14 ng Hb/ml = 2.4 µg/g	+	+	2
Haug 2011 [47]	Germany	RIDASCREEN		+	Quantitative	+	+		+	+	1
Hernandez 2014 [56]	Spain	OC-sensor		+	Quantitative	+	+		+	+	1
Hoepffner 2006 [57]	Germany	Hemocult	+						+		1
		Hb ELISA Immunodi-agnostik		+	Unknown	+		10 µg/g	+		1
Imperiale 2004 [58]	USA	Hemocult II	+						+	+	3

Imperiale 2014 [59]	Canada and USA	OC-sensor		+	Quantitative		+		+	+	1
Khalid – de Bakker 2011 [60]	Netherlands	OC-sensor		+	Quantitative	+	+		+		1
Levy 2014 [61]	USA	Inverness Clearview		+	Qualitative			50µg/g faeces	+	+	1
		Alere Clearview		+	Qualitative			6 µg/g faeces	+	+	1
		Polymedco OC-Light		+	Qualitative	+			+	+	1
		Quidel QuickVue		+	Qualitative			50µg/g faeces	+	+	1
Lieberman 2001 [12]	USA	Hemocult II	+						+	+	3
Nakama 2000 [62]	Japan	Iatro Hemcheck		+	Qualitative			unknown		+	2
Omata 2011 [63]	Japan	OC-micro		+	Quantitative	+	+		+	+	1
Park 2010 [64]	South Korea	Hemocult II	+						+	+	3
		OC-sensa		+	Quantitative	+	+		+	+	3
Sung 2003 [65]	China	Hemocult II	+						+	+	3
Wong 2014 [66]	Hong Kong	Hemosure		+	Qualitative			50ng/ml = 50 µg/g faeces	+	+	1
Wu 2014 [67]	Taiwan	ACON Laboratories		+	Qualitative			50ng/ml = 6 µg/g faeces	+	+	1

Abbreviations: ng/ml; nanogram per milliliter, µg; microgram, g; gram, AN; advanced neoplasia, CRC; colorectal cancer, gFOBT; guaiac faecal occult blood test, FIT; faecal immunochemical test.

Table 2 Overview test characteristics per study for type II studies

Study	Country	Test brand	gFOBT	FIT	FIT method	FIT 10 µg	FIT 20 µg	Other cutoff	CRC	Nr. of stools
Bouvier 1999 [68]	France	Hemocullt II	+						+	n.d.
Castiglione 2007 [69]	Italy	OC-Hemodia		+	Quantitative		+		+	1
Chiang 2014 [22]	Taiwan	OC-sensor		+	Quantitative		+		+	1
		HM-Jack			Quantitative		+		+	1
Crotta 2012 [70]	Italy	OC-sensor		+	Quantitative		+		+	1
Denters 2012 [71]	Netherlands	Hemoccult II	+						+	3
		OC-Sensor		+	Quantitative	+			+	1
Faivre 2004 [45]	France	Hemocullt II	+						+	3
Giai 2014 [72]	France	Hemocullt II	+						+	3
Itoh 1996 [73]	Japan	OC-Hemodia		+	Qualitative	+			+	1
Kronborg 1987 [74]	Denmark	Hemocullt II	+						+	3
Launoy 2005 [75]	France	Magstream		+	Quantitative			20 ng/ml (67 µg Hb/g feces)	+	2
Levi 2011 [76]	Israel	Hemocullt Sensa	+						+	3
		OC-micro		+	Quantitative			70 ng Hb/ml (14 µg Hb/g feces)	+	3
McNamara 2014 [77]	Ireland	OC-sensor		+	Quantitative		+		+	2
Nakama 1996 [78]	Japan	Monohaem		+	Qualitative				+	1
Paimela 2010 [79]	Finland	Hemocullt	+						+	3

Parente 2013 [80]	Italy	HM-Jack		+	Quantitative			100 ng/ml (250 µg Hb/g faeces)	+	1
Parra-Blanco 2010 [81]	Spain	Hemofec	+						+	3
		OC-light		+	Qualitative	+			+	1
Sieg 2002 [82]	Germany	unknown		+	Quantitative			1994-1998 10 µg/g, 1998-2000 5 µg/g	+	1
Steele 2009 [83]	Scotland	Hema-screen	+						+	3
Van Roon 2013 [84]	Netherlands	OC-Sensor		+	Quantitative	+			+	1

Abbreviations: ng/ml; nanogram per milliliter, µg; microgram, g; gram, AN; advanced neoplasia, CRC; colorectal cancer, gFOBT; guaiac faecal occult blood test, FIT; faecal immunochemical test.

Methodological quality of included studies

All characteristics of included studies can be found in Table 2.1 and 2.2.

Type I studies

The overall quality of type I studies is summarized in Figures 2.2A and 2.2B using QUADAS-II. Nineteen (83%) studies clearly included a representative spectrum of average-risk participants, which reflects population-based screening. For three studies this spectrum was unclear because the studies it were either retrospective or did not describe exclusion criteria clearly. However, all these studies were performed in an average-risk CRC screening setting and therefore included for analysis. One study (Cruz-Correa 2007) had a high risk of bias regarding the spectrum as it included patients referred for colonoscopy outside of a screening setting. However, only asymptomatic patients over the age of 55 years were included, for these reasons we choose to include this article for analysis. Eighteen (78%) of the studies had a low risk of bias concerning the index test; four were unclear because either the method of collection was not clearly described or the positivity threshold was not described. One study (4%) had a high risk of bias since the study conducted the index test differently than as advised by the manufacturer. Unanalyzable tests were only reported in five studies (22%). Twenty-two studies (96%) had low concerns regarding applicability of the index test; with one study rated as unclear because this study did not describe whether the threshold was pre-specified. All studies had a low risk concerning the reference standard with over 80% of the participants undergoing colonoscopy as the reference standard. However, many studies had missing values and FOBT-positivity rates were often lacking. The majority (87%) of studies had clear definitions of advanced adenomas; mainly defined as adenomas ≥10 mm, adenomas with at least 25% villous component, and /or high grade dysplasia.

Figure 2.2A Risk of bias and applicability concerns graph type I studies: review authors' judgements about each domain presented as percentages across included studies

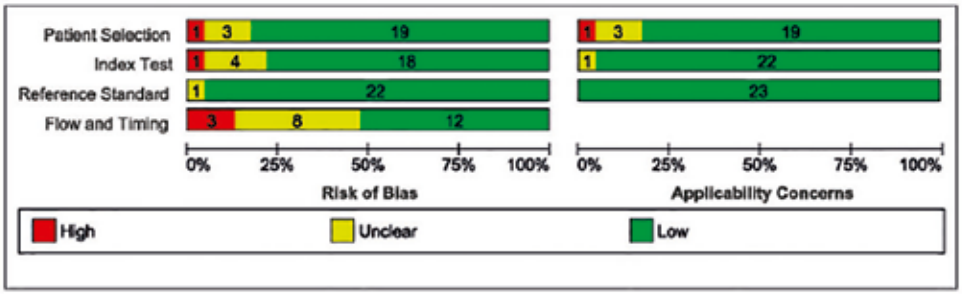


Figure 2.2B Risk of bias and applicability concerns summary type I studies: review authors' judgements about each domain for each included study

Figure 2.3A Risk of bias and applicability concerns graph type II studies: review authors' judgements about each domain presented as percentages across included studies

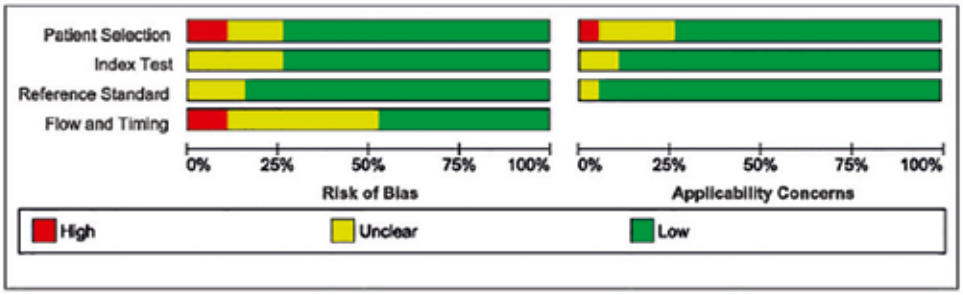
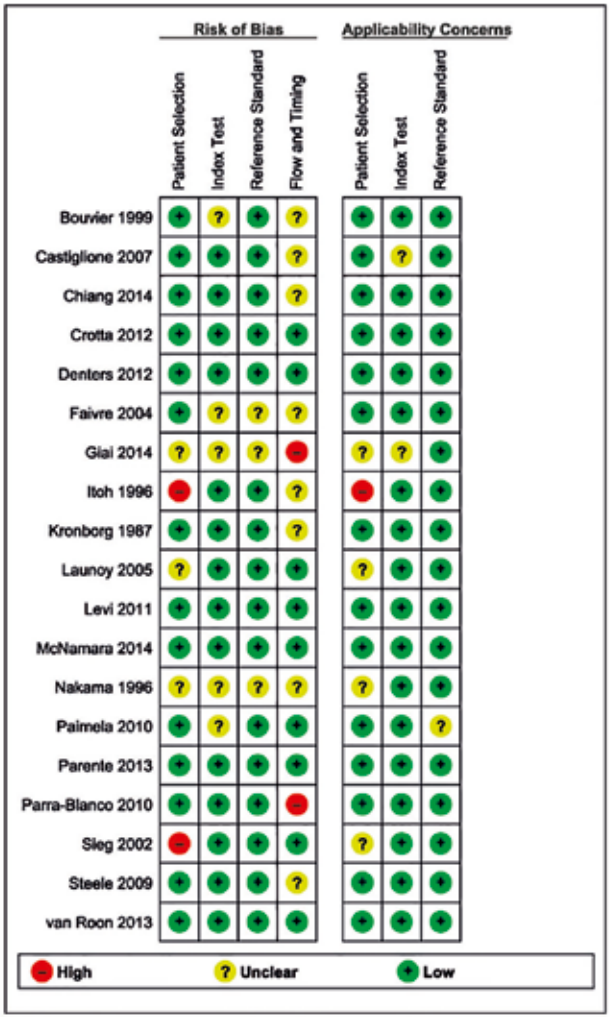


Figure 2.3B Risk of bias and applicability concerns summary type II studies: review authors' judgements about each domain for each included study



Type II studies

The overall quality of type II studies is summarized in Figures 2.3A and 2.3B. Fourteen studies (74 %) clearly included a representative spectrum of participants with an average-risk of developing advanced neoplasia. Two studies had a high risk of bias with regard to selection of patients; in one study (Itoh 1996) Japanese workers in a FIT-based screening program could have experienced gFOBT screening during earlier years. In another study (Sieg 2002), the article stated that subjects below the age of 44 could also participate if they heard of the study, but when contacted the authors stated this was not the case. Risk of bias concerning the index test was potentially present if the paper did not specify how the authors had handled non-interpretable or borderline test results. Around 84% specified their reference standard as being colonoscopy. Three studies were marked as unclear because while they used colonoscopy as a reference standard, they did not describe how many people underwent CT-colonography or DBCE in case of a failed colonoscopy. With regard to flow and timing, 47% of the studies had a low risk of bias. This was due to multiple reasons. All studies had at least 1 year of follow up, five studies had a follow-up of three years, and one study had a follow-up of four years.

Findings

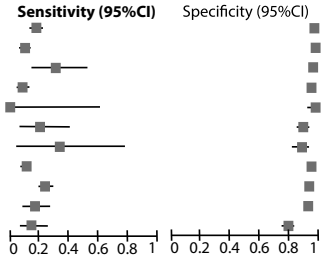
Type I studies - diagnostic test accuracy of gFOBT and FIT for advanced neoplasia

Twenty-one type I studies reported on advanced neoplasia (AN) as outcome; their median sample size was 1,046 (range 126 to 18,296). Figure 2.4 shows the Forrest plot of all included studies reporting on advanced neoplasia.

Figure 2.4 Forrest plot of all gFOBT and FIT (Type I) for advanced neoplasia
For all FIT's a cut-off of 10 mcg Hb/ g faeces was used, unless this cut-off was unavailable.

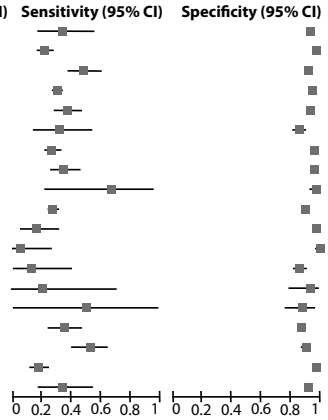
Type I gFOBT AN

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95%CI)
Ahlquist 2008	51	115	239	3359	0.18 [0.13, 0..22]	0.97 [0.96, 0.97]
Alhquist 2008	29	67	261	3407	0.10 [0.07, 0.14]	0.98 [0.98, 0.99]
Brenner 2012	8	25	18	595	0.31 [0.14, 0.52]	0.96 [0.94, 0.97]
Brenner 2013	19	92	203	1921	0.09 [0.05, 0.13]	0.95 [0.94, 0.96]
Cruz-Correa 2007	0	3	4	119	0.00 [0.00, 0.60]	0.98 [0.93, 0.99]
Graser 2009	5	26	20	225	0.20 [0.07, 0.41]	0.90 [0.85, 0.93]
Hoepffner 2006	2	17	4	133	0.33 [0.04, 0.78]	0.89 [0.82, 0.93]
Imperiale 2005	47	99	387	1972	0.11 [0.08, 0.14]	0.95 [0.94, 0.96]
Lieberman 2001	73	166	233	2413	0.24 [0.19, 0.29]	0.94 [0.93, 0.94]
Park 2010	12	49	60	639	0.17 [0.09, 0.27]	0.93 [0.91, 0.95]
Sung 2003	9	92	54	350	0.14 [0.07, 0.25]	0.79 [0.75, 0.83]



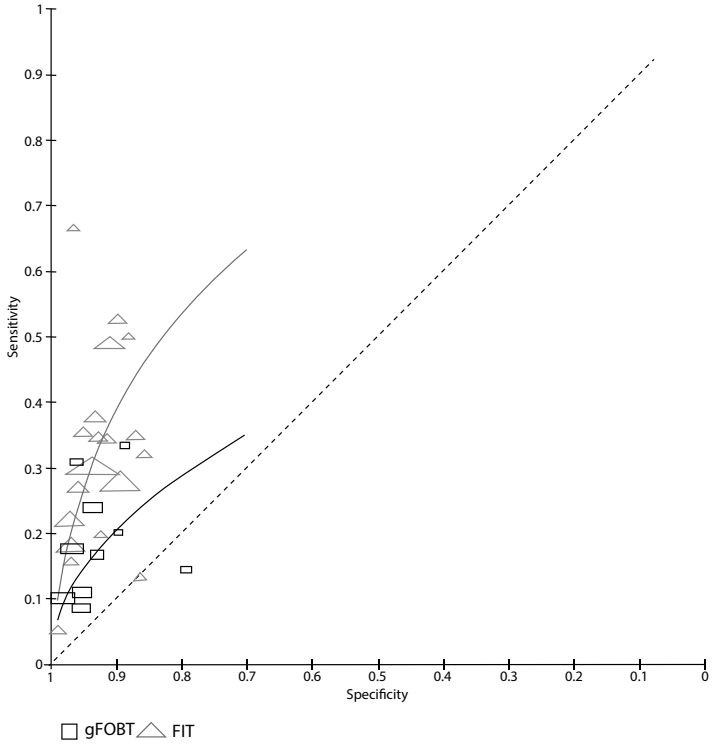
Type I FIT AN

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brenner 2012	9	45	17	575	0.35 [0.17, 0.56]	0.93 [0.90, 0.95]
Chen 2014	56	173	198	5669	0.22 [0.17, 0.28]	0.97 [0.97, 0.97]
Cheng 2002	40	448	42	4537	0.49 [0.38, 0.60]	0.91 [0.90, 0.92]
Chiu 2013	199	1131	461	16505	0.30 [0.27, 0.34]	0.94 [0.93, 0.94]
de Wijkerslooth 2012	45	76	74	1061	0.38 [0.29, 0.47]	0.93 [0.92, 0.95]
Graser 2009	8	37	17	223	0.32 [0.15, 0.54]	0.96 [0.81, 0.90]
Haug 2011	66	87	177	1995	0.27 [0.22, 0.33]	0.96 [0.95, 0.97]
Hernandez 2014	34	33	62	650	0.35 [0.26, 0.46]	0.95 [0.93, 0.97]
Hoepffner 2006	4	5	2	145	0.67 [0.22, 0.96]	0.97 [0.92, 0.99]
Imperiale 2014	228	964	594	8203	0.28 [0.25, 0.31]	0.89 [0.89, 0.90]
Khalid-de Bakker 2011	6	9	32	282	0.16 [0.06, 0.31]	0.97 [0.94, 0.99]
Levy 2014	1	2	18	196	0.05 [0.00, 0.26]	0.99 [0.96, 1.00]
Levy 2014	2	40	13	253	0.13 [0.02, 0.40]	0.86 [0.82, 0.90]
Levy 2014	1	3	4	36	0.20 [0.01, 0.72]	0.92 [0.79, 0.98]
Levy 2014	1	6	1	44	0.50 [0.01, 0.99]	0.88 [0.76, 0.95]
Omata 2011	27	131	50	877	0.35 [0.25, 0.47]	0.87 [0.85, 0.89]
Park 2010	38	71	34	627	0.53 [0.41, 0.65]	0.90 [0.87, 0.92]
Wong 2014	30	120	137	3840	0.18 [0.12, 0.25]	0.97 [0.96, 0.97]
Wu 2014	10	83	19	895	0.34 [0.18, 0.54]	0.92 [0.90, 0.93]



Sensitivities for detecting AN ranged from 0% to 33% for gFOBT and from 5% to 67% for FIT screening. Specificities ranged from 79% to 98% and from 86% to 99%, respectively. The cut-off for positivity of FIT varied between 2.4 to 50 µg Hb/g faeces. The summary curve estimated by the HSROC model for all Type I studies for AN can be found in Figure 2.5. FIT showed a higher discriminative ability for AN than gFOBT (p=0.002).

Figure 2.5 Summary curve using the HSROC model for gFOBT and FIT for multiple cut-offs for advanced neoplasia (Type I) Scale of individual study points is based on sample size.

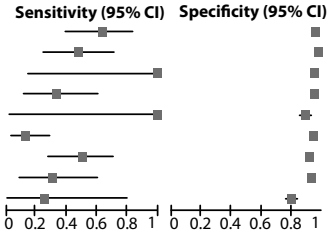


In addition, sensitivities and specificities were calculated solely for those studies reporting on FIT screening with a cut-off value of 10 µg Hb/g faeces and 20 µg Hb/g. Analyses for cut-off 10 µg Hb/g faeces contained both qualitative as quantitative FITs. The sensitivity of FIT screening for detection of AN ranged between 5% and 67% with a cut-off of 10 µg Hb/g, and from 13 to 44% with a cut-off of 20 µg Hb/g. The sensitivity for AN was lower for gFOBT screening with a summary sensitivity of 16% (95% CI 12-21%) compared to 31% 995% CI 25-39% for FIT with a cut-off of 10 µg Hb/g and 27% (95% CI 21-34%) with a cut-off of 20 µg Hb/g Specificities of FIT screening for detecting AN ranged from 87% to 97% for a cut-off of 10 µg Hb/g, and from 89% to 100% for a cut-off of 20 µg Hb/g. No significant differences in summary specificity for AN were found between gFOBT (94%; 95% CI 92-96%), FIT with a cut-off of 10 µg Hb/g (95%; 95% CI 92-97%) and with a cut-off of 20 µg Hb/g (97%; 95% CI 94-98%).

Figure 2.6 Forrest plot of all gFOBT and FIT (Type I) for colorectal cancer For all FIT's a cut-off of 10 µg Hb/ g faeces was used, unless this cut-off was unavailable.

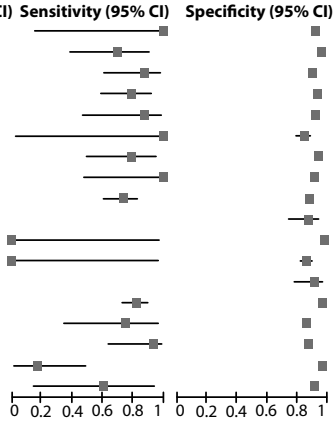
Type I gFOBT CRC

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahlquist 2008	12	154	7	3591	0.63 [0.38, 0.84]	0.96 [0.95, 0.97]
Ahlquist 2008	9	87	10	3658	0.47 [0.24, 0.71]	0.98 [0.97, 0.98]
Brenner 2012	2	31	0	613	1.00 [0.16, 1.00]	0.95 [0.93, 0.97]
Brenner 2013	5	106	10	2114	0.33 [0.12, 0.62]	0.95 [0.94, 0.96]
Graser 2009	1	30	0	245	1.00 [0.03, 1.00]	0.89 [0.85, 0.93]
Imperiale 2004	4	142	27	2332	0.13 [0.04, 0.30]	0.94 [0.93, 0.95]
Liebermann 2001	12	227	12	2634	0.50 [0.29, 0.71]	0.92 [0.91, 0.93]
Park 2010	4	57	9	690	0.31 [0.09, 0.61]	0.92 [0.90, 0.94]
Sung 2003	1	100	3	401	0.25 [0.01 0.81]	0.80 [0.76, 0.83]



Type I FIT CRC

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brenner 2012	2	52	0	592	1.00 [0.16, 1.00]	0.92 [0.90, 0.94]
Chen 2014	9	220	4	5863	0.69 [0.39, 0.91]	0.96 [0.96, 0.97]
Cheng 2002	14	474	2	4577	0.88 [0.62, 0.98]	0.91 [0.90, 0.91]
Chiu 2013	22	1308	6	16960	0.79 [0.59, 0.92]	0.93 [0.92, 0.93]
de Wijkerslooth 2012	7	114	1	1134	0.88 [0.47, 1.00]	0.91 [0.89, 0.92]
Graser 2009	1	44	0	240	1.00 [0.03, 1.00]	0.85 [0.80, 0.89]
Haug 2011	11	142	3	2169	0.79 [0.49, 0.95]	0.94 [0.93, 0.95]
Hernandez 2014	5	62	0	712	1.00 [0.48, 1.00]	0.92 [0.90, 0.94]
Imperiale 2014	48	1144	17	8780	0.74 [0.61, 0.84]	0.88 [0.88, 0.89]
Levy 2014	0	7	0	45	Not estimable	0.87 [0.74, 0.94]
Levy 2014	0	3	1	213	0.00 [0.00, 0.97]	0.99 [0.96, 1.00]
Levy 2014	0	42	1	265	0.00 [0.00, 0.97]	0.86 [0.82, 0.90]
Levy 2014	0	4	0	40	Not estimable	0.91 [0.78, 0.97]
Nakama 2000	79	753	17	16815	0.82 [0.73, 0.89]	0.96 [0.95, 0.96]
Omata 2011	6	151	2	926	0.75 [0.35, 0.97]	0.86 [0.84, 0.88]
Park 2010	12	97	1	660	0.92 [0.64, 1.00]	0.87 [0.85, 0.89]
Wong 2014	2	148	10	3967	0.17 [0.02, 0.48]	0.96 [0.96, 0.97]
Wu 2014	3	90	2	912	0.60 [0.15, 0.95]	0.91 [0.89, 0.93]

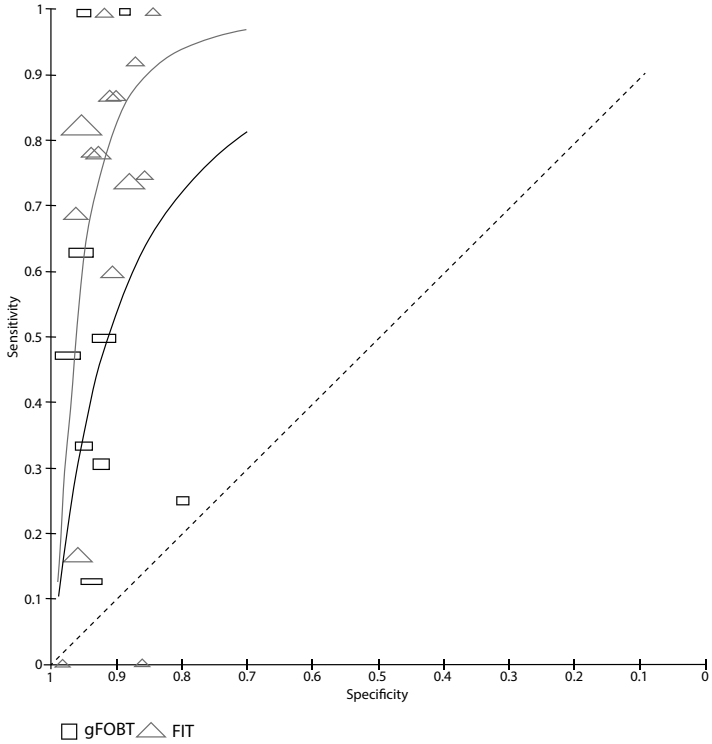


Type I studies - diagnostic test accuracy of gFOBT and FIT for colorectal cancer

Nineteen type I studies reported on colorectal cancer (CRC) as separate outcome measure; their median sample size was 2,235 (range 285 to 18,269). Figure 2.6 shows the Forrest plot of all included studies reporting on colorectal cancer.

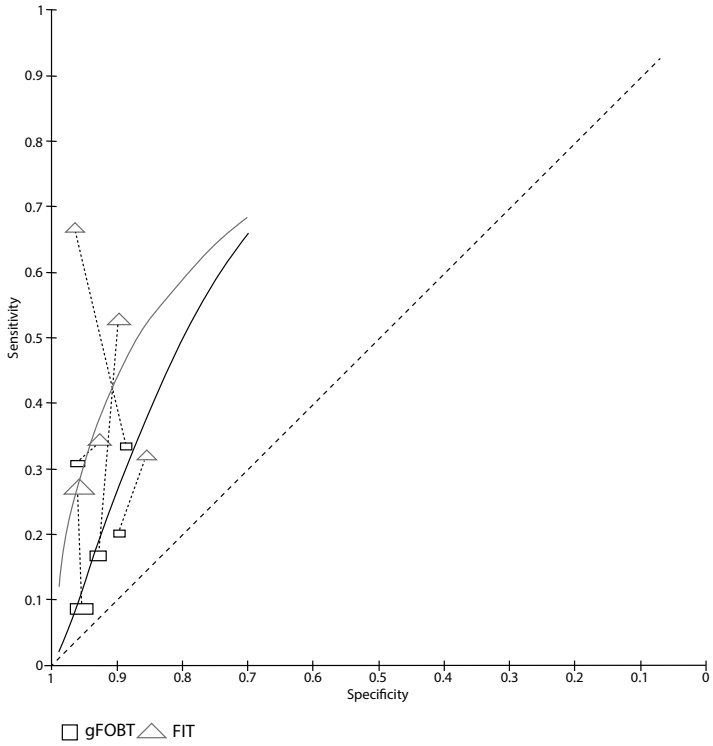
Sensitivities ranged from 13% to 100% for gFOBT, and from 0% to 100% for FIT. Specificities ranged from 80% to 98% for gFOBT and from 85% to 96% for FIT. The cut-off for positivity of FIT varied between 2.4 to 50 µg Hb/g faeces. The summary curve estimated by the HSROC model for AN can be found in Figure 2.7. FIT showed a higher discriminative ability for AN than gFOBT (p=0.025).

Figure 2.7 Summary curve type I CRC. Summary curve using the HSROC model for gFOBT and FIT (Type I) for multiple cut-offs for colorectal cancer. Scale of individual study points is based on sample size.



In addition, sensitivities and specificities were calculated solely for those studies reporting on a cut-off of 10 µg Hb/g faeces and 20 µg Hb/g for FIT. Sensitivities for CRC ranged from 13% to 100% for gFOBT, from 0% to 100% for a FIT cut-off of 10 µg Hb/g, and from 50% to 100% for a cut-off of 20 µg Hb/g. Sensitivity for CRC was lower for gFOBT with a summary sensitivity of 41% (95% CI 29-54%), compared to 81% (95% CI 71-89%) for a FIT cut-off of 10 µg Hb/g, and 76% (95% CI 61-86%) for a FIT cut-off of 20 µg Hb/g. Specificities for FIT ranged from 87% to 99% when using a cut-off of 10 µg Hb/g, and from 88% to 96% with a cut-off of 20 µg Hb/g. No significant differences in summary specificity for colorectal cancer were found between gFOBT (94%; 95% CI 91-95%), FIT with a cut-off of 10 µg Hb/g (93%; 95% CI 91-95%), and with a cut-off of 20 µg Hb/g (93%; 95% CI 90-95%).

Figure 2.8 Linked-HSROC curve of studies (Type I) with outcome advanced neoplasia (including: Brenner 2012, Graser 2009, Brenner 2013 and Haug 2011 combined, Park 2010, Hoepffner 2006). Scale of individual study points is based on sample size.



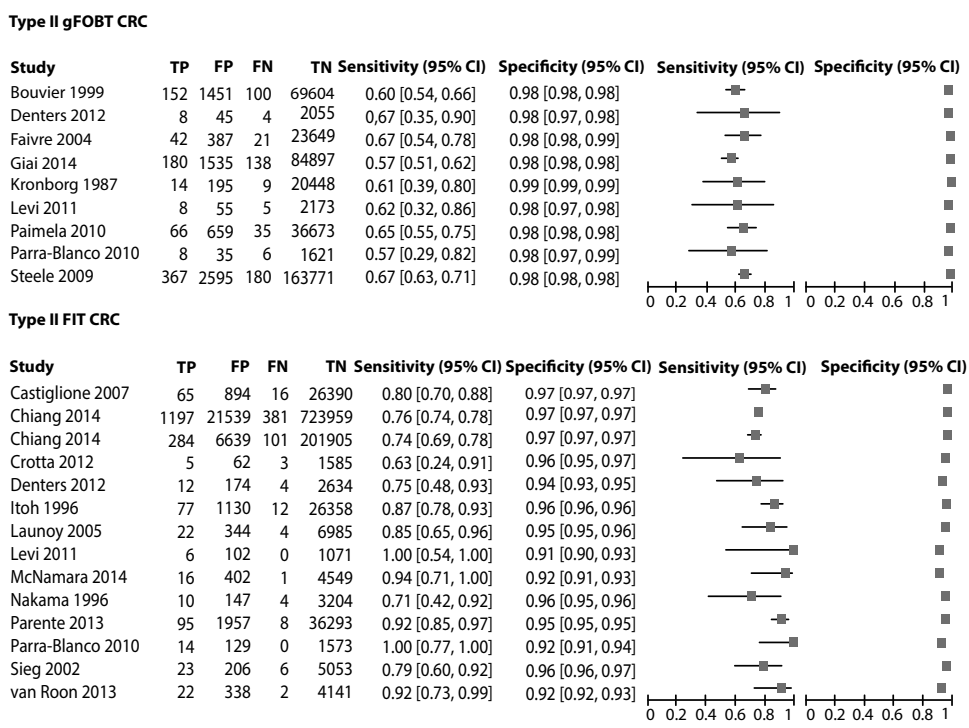
Type I studies – linked ROC

Four studies (Brenner 2010, Brenner 2013 & Haug 2011 combined, Graser, 2009, Hoepffner 2006, Park 2010) compared FIT and gFOBT in the same population. The cut-off for positivity of FIT varied between 2.4 to 10 µg Hb/g feces. The summary curve estimated by the HSROC model for linked Type I studies for AN, can be found in Figure 2.8. FIT showed a higher discriminative ability for AN than gFOBT (p=0.073).

Type I studies heterogeneity analyses

There was a significant difference in sensitivity or specificity, or both, for males versus females for FIT, both for outcome AN as CRC (p<0.001). For gFOBT, difference in accuracy for males versus females was significant for outcome AN but not for CRC (p=0.002 and p=0.638, respectively). There was no evidence (all p-values >0.01) to suggest a difference in sensitivity or specificity, or both, between studies using one, two or three stools per screening round. There was

Figure 2.9 Forrest plot of all gFOBT and FIT (Type II) for colorectal cancer. For all FIT's a cut-off of 10 mcg Hb/g faeces was used, unless this cut-off was unavailable.



no significant difference in sensitivity or specificity, or both, between studies using a quantitative or a qualitative FIT at a cut-off of 10 µg Hb/g for the outcome AN as well as CRC (p = 0.645 and p = 0.216, respectively).

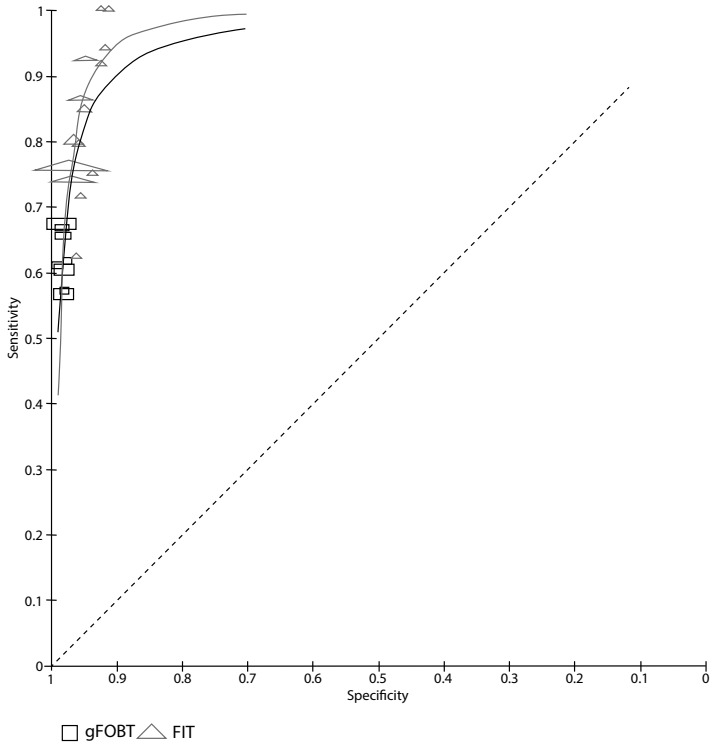
Type I studies sensitivity analyses

For the analyses including all cut-offs, a sensitivity analysis was undertaken by excluding the studies that yielded an high risk of bias following the QUADAS assessment (Alhquist 2008, Brenner 2013, Cruz-Correa 2007, Hoepffner 2006, Omata 2011, Sung 2013). Even when excluding these studies, FIT remained significantly superior to gFOBT in the HSROC model.

Type II studies - diagnostic test accuracy of gFOBT and FIT for colorectal cancer

There were nineteen type II studies that reported on CRC as separate outcome; their median sample size was 7,355 (range 1,179 to 747,076). Figure 2.9 shows the Forrest plot of all included studies reporting on CRC.

Figure 2.10 Summary curve type II studies. Summary curve using the HSROC model for gFOBT and FIT (Type II) for multiple cut-offs for colorectal cancer. Scale of individual study points is based on sample size.



Sensitivities for detection of CRC ranged from 57% to 67% for gFOBT, and from 63% to 100% for FIT. Specificities ranged from 98% to 99% for gFOBT and from 91% to 97% for FIT. The cut-off for positivity of FIT varied between 5 to 250 µg Hb/g faeces. The summary curve estimated by the HSROC model, for all Type II studies for CRC, can be found in Figure 2.10. FIT showed a higher discriminative ability for FIT than gFOBT ($p<0.001$)

In addition, sensitivities and specificities were calculated separately for those studies reporting on a FIT cut-off of 10 µg Hb/g faeces, and those with a cut-off of 20 µg Hb/g for CRC. Sensitivities for detection of CRC ranged from 75% to 100% when using a cut-off of 10 µg Hb/g , and from 63% to 94% with a cut-off of 20 µg Hb/g. Sensitivity for CRC was lower for gFOBT with a summary sensitivity of 63% (95% CI 58-67%), compared to 87% (95% CI 80-92%) for a cut-off of 10 µg Hb/g, and 88% (95% CI 74-94%) for a cut-off of 20 µg Hb/g. Specificities for FIT ranged between studies from 92% to 96% when using a cut-off of 10 µg

Hb/g, and from 92% to 97% for a cut-off of 20 µg Hb/g. Specificity for CRC was higher for gFOBT with a summary specificity for colorectal cancer of 98% (95% CI 98-99%) for gFOBT, compared to 92% (95% CI 92-95%) for FIT with a cut-off of 10 µg Hb/g, and 95% (95% CI 93-97%) at a cut-off of 20 µg Hb/g.

Type II studies heterogeneity analyses

There was a significant difference in sensitivity or specificity, or both, for males versus females for gFOBT, for the outcome CRC ($p<0.001$). There was no significant difference in sensitivity or specificity, or both, between studies using a quantitative or a qualitative FIT at a cut-off of 10 µg Hb/g for the outcome CRC ($p = 0.684$). Heterogeneity related to the number of stools per screening round was not performed for gFOBT since all studies used three stools. For the following covariates analyses were not possible due to convergence difficulties; gender for FIT, number of stools for FIT.

Type II studies sensitivity analyses

For the analyses including all cut-offs, a sensitivity analysis was undertaken by excluding the studies that yielded an high risk of bias following the QUADAS assessment (Giai 2014, Itoh 1996 Nakama 1996, Parra-Blanco 2010, Sieg 2002).[72 73 78 81 82] Even when excluding these studies, FIT remained significantly superior to gFOBT in the HSROC model. The effect of removing studies, in which the percentage of participants with a positive FOBT that underwent the reference standard was unknown, was herein evaluated (Nakama 1996, Giai 2014).[72 78]

Summary of Findings

Diagnostic accuracy of gFOBT compared to FIT					
Patients/ population	Asymptomatic, average-risk individuals over the age of 40 years undergoing colorectal cancer screening				
Prior testing	Only the results of the first screening round were included in this analysis				
Settings	Population- based colorectal cancer screening				
Index test	Guaiac faecal occult blood test or faecal immunochemical test				
Importance	Many screening programmes worldwide are currently changing from gFOBT to FIT-based screening				
Reference standard	Colonoscopy is the gold standard for the diagnosis of colorectal cancer which was used as the reference standard. Only in case a colonoscopy was not complete a CT-colonography (or double contrast barium enema) was used as a surrogate.				
Studies	Prospective and retrospective studies including average-risk individuals invited for colorectal cancer screening Type I: all screenees underwent both the index test and colonoscopy (n=23). Type II: only screenees with a positive index test underwent colonoscopy and all screen negatives were followed for at least one year (n=19).				
Quality concerns	Due to strict inclusion criteria most studies were of high quality. Few studies had unclear risk of bias due to poor reporting of a pre-specified cut-off value. Only three studies had a high risk of bias regarding the selection of study population. Regarding these studies, sensitivity analyses showed significant differences in outcome when excluding these studies from analyses.				
Test /subgroup*					
	studies (participants)	summary sensitivity	summary specificity	summary sensitivity	summary specificity
	gFOBT/FIT	gFOBT	gFOBT	FIT*	FIT*
		(%, 95% CI)	(%, 95% CI)	(%, 95% CI)	(%, 95% CI)
Type I					
advanced neoplasia	10 (15.741) / 16 (55.881)	15 (12-19)	94 (91-96)	31 (25-39)	95 (92-97)
colorectal cancer	8 (15.465) / -15 (69.998)	41 (29-54)	94 (91-96)	82 (71-89)	93 (90-95)
Type II					
colorectal cancer	9 (413.191) / 14 (1.082.153)	63 (58-67)	98 (98-99)	87 (80-92)	92 (92-95)
Conclusions	**The results of this systematic review concludes that FIT is the preferred tool for FOBT-based population CRC screening due to the higher sensitivity and comparable similar specificity as compared to gFOBT.**				

CAUTION: The results on this table should not be interpreted in isolation from the results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.
* results for FIT cut-off 10 µg Hb/g faeces are shown

DISCUSSION

Summary of main results

The main results are presented in the Summary of Findings. For this review we chose to include two types of studies and report the results separately as they differ in type of study and yield comparable but not similar results. We included twenty-three type I studies and nineteen type II studies. Four type I studies evaluated the diagnostic accuracy of gFOBT, thirteen studies evaluated FIT and four studies assessed both gFOBT and FIT tests. Twenty-one studies evaluated the diagnostic accuracy for advanced neoplasia and nineteen studies evaluated the diagnostic accuracy for colorectal cancer. FIT showed a higher discriminative ability than gFOBT as assessed by the HSROC curve for both advanced neoplasia ($p=0.002$) and CRC ($p=0.025$). As type I studies allowed the use of multiple FIT cut-offs within one study population, the two most commonly used cut-offs worldwide (10 and 20 µg Hb/g faeces) were analysed separately and these results were in line with the overall analyses. Nineteen type II studies were included, with six studies evaluating the diagnostic accuracy of gFOBT, ten studies evaluating FIT, and three studies combining both. In type II studies FIT also showed a higher discriminative ability than gFOBT as assessed by the HSROC curve for CRC ($p<0.001$).

The results of this systematic review demonstrate that FIT is the preferred tool for FOBT-based population screening due to the superior sensitivity and similar specificity compared as to gFOBT screening. Furthermore, beside some qualitative tests, many FITs are quantitative tests which allows use of different cut-offs to tailor for screening resources and colonoscopy capacity. Finally, various studies have consistently shown that FIT screening is associated with higher uptake than gFOBT screening which is an important finding to reach a high coverage of the target population (i.e. cumulative uptake).

Strengths

The results of this review are based on strict and thorough searching without any language or date restrictions. The use of diagnostic test accuracy, or randomised controlled trial filters may lead to the loss of some studies, for this reason we have not used any filters.[85] Two independent reviewers identified and extracted data from the studies, thus decreasing inaccuracies related to single-person data extraction.[86] All included studies reported the results for average-risk, asymptomatic individuals after the age of 40 years, making our results reflective of a screening population. Also, data for different cut-offs were retrieved, in cases where this had not already been reported in the original publication, contacting

the author provided additional data for most studies allowing sub-analyses with the two most commonly used cut-offs. These cut-offs were converted to the internationally used measuring standard of µg Hb/g faeces.[35] To avoid potential bias caused by the use of an inappropriate reference standard, e.g. barium-enema or sigmoidoscopy, we restricted the studies to those with colonoscopy as the reference standard. As mentioned above, two types of studies were included for this review and analysed separately. The inclusion of type I studies allowed evaluating both advanced neoplasia and colorectal cancer as outcome of FOBT-based screening. Advanced neoplasia is of special importance because by removing adenomas development of colorectal cancer and CRC deaths might be prevented.[29 87] In type II studies sensitivities and specificities were calculated with the use of interval carcinomas identified through adequate follow-up as a surrogate for the gold standard; colonoscopy. This different character of type I and type II studies may explain the observed differences in sensitivity and specificity. The use of interval carcinomas as endpoint in type II studies may underestimate the true proportion of false negatives, as by definition only those cancers were reported that had become clinically evident during the observation period. On the other hand, Type II studies are more reflective of a FOBT based screening programme in a general population. In these studies willingness to undergo FOBT as a primary screening tool was assessed whereas in type I studies participants had to be willing to undergo a full colonoscopy irrespective of the FOBT-result. For this reason type II studies are also often performed in larger populations. Combining both types of studies provides insight on both settings, and results in a broad evaluation of FOBT diagnostic test accuracy in colorectal cancer screening. Type I studies give insight in test sensitivity, whereas type II studies give insight in program sensitivity. The overall quality of included studies was high, supporting the validity of the results of our analyses.

Weaknesses

This systematic review was designed to evaluate the diagnostic test accuracy of two types of FOBTs commonly used for colorectal cancer screening. Even though diagnostic test accuracy is of major importance in screening, usability of the test and participation, e.g. willingness to undergo the screening test, is also very relevant. One major limitation of this review is that these latter points have not been taken into account for this review as they do not involve diagnostic test accuracy. Yet, these factors are also of importance when estimating screening efficacy on population level (programme sensitivity and specificity). The ultimate purpose of screening programmes is a decrease in colorectal cancer-related mortality. However, diagnostic test accuracy can only be used as a surrogate in estimating mortality decrease after screening. In past years, results of large

prospective gFOBT-based screening trials have been published that unfortunately could not be included in this review as their main outcome was mortality. Mortality rates could not be converted into contingency tables to calculate sensitivity and specificity. We excluded non-interpretable test results and FOBT-positives who refused to undergo the reference standard from the 2x2 table and in consequence from the meta-analysis. Missing FOBT results are likely to be completely random (incidental missing data) and will not lead to biased estimates of test accuracy. Because only the participants who received the reference standard were included in the type I studies analysis (complete case analysis) and positive participants who did not receive the reference standard in the type II studies were excluded from analysis, estimates of the accuracy of the diagnostic tests could be biased.[88] The percentage of participants that did not undergo colonoscopy is reported in the characteristics of included studies table. Because of the large size of the studies, we believe that excluding these participants for whom the reference standard result is missing is mostly at random and will not bias the results but will only decrease precision. We attempted to conduct a comprehensive search for studies, but the fact that the studies awaiting classification have not yet been incorporated may be a source of potential bias.

Another limitation is the inability to explore the sources of heterogeneity concerning age, gender, ethnicity, adenoma type, tumour localization and stage distribution, because of limited information in the included studies. This problem was more prominent in type II studies, which often described large populations without prospective registration of study outcomes, but rather national databases of hospitals or cancer registries. The chosen random effect model accounts to a certain level for heterogeneity across studies by considering both within and between-study variation. This leads to a greater precision of the pooled estimates, but larger confidence.[89] For only for a limited amount of studies was a direct comparison possible, as most studies did not perform both tests in the same patients. This could be a limitation, as results from non-comparative studies may differ from comparative studies.[90] However, in this review results from the comparative studies are in line with the overall results. Many test brands are available, and sub analyses of these brands was not possible due to limited data and the use of different sub-types of the same brand.

A previously published meta-analysis evaluated the test accuracy of FIT for colorectal cancer but not advanced neoplasia and did not compare these results to the performance of gFOBT.[91] To the best of our knowledge there is only one other systematic review comparing the diagnostic test accuracy of gFOBT and FIT.[75] The highest summary sensitivity of this study for FIT (OC-sensor) was 87%

and specificity was 93%, and for gFOBT (Hemoccult,) the summary sensitivity was 47% and summary specificity was 95%. They did not distinguish between type I, type II, and case-control studies, possibly leading to underestimation (type II studies) or overestimation (case-control) studies of test accuracy.[30] Even though this large variation in type of studies included, their search strategy was very limited yielding only 761 hits and 22 inclusions. Finally, only three brands of FOBTs were included in their meta-analysis. In general this is the first review to adequately and systematically present an evaluation of gFOBT and FIT screening in an average-risk population.

Applicability of findings to the review question

All participants included in this review were asymptomatic, average-risk individuals over the age of 40 years old, and invited for colorectal cancer screening, making the findings of this review extremely relevant for colorectal cancer screening programmes. Two types of studies were included. Type I studies are more homogenous than type II studies, yet may be less representative of a FOBT-based screening population. This is due to the fact that all screenees had to be willing to undergo colonoscopy. For type II studies false negatives were identified through interval carcinomas that were identified during follow-up. This might give an underestimation of test sensitivity, yet is appropriately representative of FOBT population-based screening programmes.

Implications for practice

Faecal immunochemical testing has a superior sensitivity compared to guaiac faecal occult blood testing and is the preferred method of occult blood screening in terms of diagnostic test accuracy. Test usability and participation were not evaluated for the purpose of this review. The summary of findings table should be interpreted with acknowledgement of this. However, FIT's quantitative nature allows the use of different cut-offs tailoring to screening resources and colonoscopy capacity[24]. Furthermore, it should be noted that several studies consistently reported higher rates of participation for FIT than for gFOBT screening.[92 93] Both gFOBT and FIT have lower sensitivity for colorectal cancer than colonoscopy as gold standard. However, when combining test accuracy with participation FIT-based screening in many populations results in a higher diagnostic yield of advanced neoplasia compared to other CRC screening methods.[10 28]

Implications for research

Future studies should be conducted in a prospective manner mimicking population-based colorectal cancer screening and targeting average-risk populations. We encourage authors to systematically report data on participation, positivity rate and colonoscopy adherence. Also, future studies should report a clear definition

of advanced neoplasia and interval carcinomas. In the included studies definitions of interval carcinomas were often vague or completely missing. The ultimate purpose of colorectal cancer screening is decreasing mortality, so future studies should be conducted to compare long-term follow up on mortality between gFOBT- and FIT-based colorectal cancer screening programmes.

REFERENCES

1. Andermann A, Blancquaert I, Beauchamp S, et al. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bull World Health Organ 2008;**86**:317-19.
2. Wilson JM, Jungner YG. Principles and practice of mass screening for disease. Public Health Papers 34 1968;**65**:281-393.
3. Ciatto S, Martinelli F, Castiglione G, et al. Association of FOBT-assessed faecal Hb content with colonic lesions detected in the Florence screening programme. Br J Cancer 2007;**96**:218-21.
4. Edwards JB. Screening for colorectal cancer using faecal blood testing: varying the positive cut-off value. Pathology 2005;**37**:565-68.
5. Levi Z, Rozen P, Hazazi R, et al. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 2007;**146**:244-55.
6. Rozen P, Waked A, Vilkin A, et al. Evaluation of a desk top instrument for the automated development and immunochemical quantification of fecal occult blood. Med Sci Monit 2006;**12**:MT27-MT32.
7. van Doorn SC, Stegeman I, Stroobants AK, et al. Fecal immunochemical testing results and characteristics of colonic lesions. Endoscopy 2015;**47**(11):1011-7.
8. van Rijn JC, Reitsma JB, Stoker J, et al. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol 2006;**101**:343-50.
9. Young GP. Population-based screening for colorectal cancer: Australian research and implementation. J Gastroenterol Hepatol 2009;**24 Suppl 3**:S33-S42.
10. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. Nat Rev Dis Primers 2015;**1**:15065.
11. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Studyh. N Engl J Med 1993;**328**:1365-71.
12. Lieberman DA, Weiss DG, Veterans Affairs Cooperative Study G. One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon. New England Journal of Medicine 2001;**345**(8):555-60.
13. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;**348**:1467-71.
14. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;**348**:1472-77.
15. Lindholm E, Brevinge H, Haglind E. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. Br J Surg 2008;**95**(8):1029-36.
16. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. The American journal of gastroenterology 2008;**103**(6):1541-9.
17. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;**369**(13):1106-14.
18. Allison JE, Tekawa IS, Ransom LJ, et al. A comparison of fecal occult-blood tests for colorectal-cancer screening. N Engl J Med 1996;**334**:155-59.
19. Young GP, St John DJ, Winawer SJ, et al. Choice of fecal occult blood tests for colorectal cancer screening: recommendations based on performance characteristics in population studies: a WHO (World Health Organization) and OMED (World Organization for Digestive

Endoscopy) report. Am J Gastroenterol 2002;**97**:2499-507.

20. Brenner H, Tao S, Haug U. Low-dose aspirin use and performance of immunohcmical fecal occult blood tests. JAMA 2010;**304**(22):2513-20.

21. Levi Z, Rozen P, Hazazi R, et al. Sensitivity, but not specificity, of a quantitative immunochemical fecal occult blood test for neoplasia is slightly increased by the use of low-dose aspirin, NSAIDs, and anticoagulants. Am J Gastroenterol 2009;**104**:933-38.

22. Chiang TH, Chuang SL, Chen SL, 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**:1317-26.

23. Hundt S, Haug U, Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. Ann Intern Med 2009;**150**(3):162-9.

24. Wilschut JA, Habbema JD, van Leerdam ME, et al. Fecal occult blood testing when colonoscopy capacity is limited. J Natl Cancer Inst 2011;**103**(23):1741-51.

25. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut 2015;**64**(5):784-90.

26. Schreuders EH, Grobbee EJ, Spaander MCW, et al. Advances in fecal tests for colorectal cancer screening. Curr Treat Options Gastroenterol 2016;**14**(1):152-62.

27. Spada C, Barbaro F, Andrisani G, et al. Colon capsule endoscopy: What we know and what we would like to know. World J Gastroenterol 2014;**20**(45):16948-55.

28. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.

29. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 1997;**112**:594-642.

30. Deeks JJ, Bossuyt PM, Gatsonis C. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0.0*: The Cochrane Collaboration, 2009.

31. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. J Clin Epidemiol 2009;**62**(1):5-12.

32. Lansdorp-Vogelaar I, von Karsa L, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Introduction. Endoscopy 2012;**44 Suppl 3**(Suppl 3):SE15-30.

33. 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.

34. van Roon AHC, van Dam L, Zauber AG, et al. Protocol: Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals. . Cochrane Database of Systematic Reviews 2011(8).

35. Fraser CG, Allison JE, Halloran SP, et al. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. Journal of the National Cancer Institute 2012;**104**(11):810-4.

36. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic

accuracy studies. Ann Intern Med 2011;**155**(8):529-36.

37. Macaskill P, Gatsonis C, Deeks JJ, et al. Chapter 10: Analysing and Presenting Results. In: Deeks Jj BPMGC, ed. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1 0: The Cochrane Collaboration. Available from: http://srdta.cochrane.org/, 2010.

38. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;**20**:2865-84-65-84.

39. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002;**21**(4):589-624.

40. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis ofsensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;**58**:982-90-82-90.

41. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;**327**(7414):557-60.

42. Leeflang MM, Deeks JJ, Gatsonis C, et al. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008;**149**:889-97.

43. Song F, Khan KS, Dinnes J, et al. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. Int J Epidemiol 2002;**31**:88-95.

44. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;**58**:882-93.

45. Faivre J, Dancourt V, Lejeune C, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology 2004;**126**(7):1674-80.

46. 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.

47. Haug U, Kuntz KM, Knudsen AB, et al. Sensitivity of immunochemical faecal occult blood testing for detecting left-vs right-sided colorectal neoplasia. Br J Cancer 2011;**104**(11):1779-85.

48. Ahlquist DA, Sargent DJ, Loprinzi CL, et al. Stool DNA and occult blood testing for screen detection of colorectal neoplasia. Ann Intern Med 2008;**149**(7):441-50, W81.

49. Brenner G, Faure H, Heuer S, et al. Detection of colorectal findings for cancer prevention by immunochemical stool test with different sensitivity levels. Zeitschrift fur gastroenterologie 2012;**50**(10):1083-88.

50. Chen, Y Y, Chen TH, et al. Accuracy of immunochemical fecal occult blood test for detecting colorectal neoplasms in individuals undergoing health check-ups. Advances in Digestive Medicine 2014;**1**:74-79.

51. Cheng, Wong, Hong, et al. Colorectal cancer screening in asymptomatic adults: Comparison of colonoscopy, sigmoidoscopy and fecal occult blood tests. J Formos Med Assoc 2002;**101**(10):685-90.

52. Chiu, Lee, Tu, et al. Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clin Gastroenterol Hepatol 2013;**11**(7):832-38.

53. Cruz-Correa M, Schultz K, Jagannath S, et al. Performance characteristics and comparison of two fecal occult blood tests in patients undergoing colonoscopy.

Dig Dis Sci 2007;**52**(4):1009-13.

54. De W, Stoop, t Bossuy, et al. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced Neoplasia. Am J Gastroenterol 2012;**107**(10):1570-78.

55. Graser A, Stieber P, Nagel D, et al. Comparison of CT colonography, colonoscopy, sigmoidoscopy and faecal occult blood tests for the detection of advanced adenoma in an average risk population. Gut 2009;**58**(2):241-8.

56. Hernandez V, Cubiella J, Gonzalez-Mao MC, et al. Fecal immunochemical test accuracy in average-risk colorectal cancer screening. World J Gastroenterol 2014;**20**(4):1038-47.

57. Hoepffner N, Shastri YM, Hanisch E, et al. Comparative evaluation of a new bedside faecal occult blood test in a prospective multicentre study. Aliment Pharmacol Ther 2005;**23**:145-54.

58. Imperiale, Ransohoff, Itzkowitz, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;**351**(26):2704-14.

59. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. New Engl J Med 2014;**370**:1287-97.

60. Khalid-de Bakker CA, Jonkers DM, Sanduleanu S, et al. Test performance of immunologic fecal occult blood testing and sigmoidoscopy compared with primary colonscopy screening for colorectal advanced adenomas. Cancer Prev Res 2011;**4**:1563-71.

61. Levy BT, Bay C, Xu Y, et al. Test characteristics of faecal immunochemical tests (FIT) compared with optical colonoscopy. J Med Screen 2014;**21**(3):133-43.

62. Nakama H, Zhang B, Fattah AS, et al. Colorectal cancer in iron deficiency anemia with a positive result on immunochemical fecal occult blood. Int J Colorectal Dis 2000;**15**(5-6):271-4.

63. Omata F, Shintani A, Isozaki M, et al. Diagnostic performance of quantitative fecal immunochemical test and multivariate prediction model for colorectal neoplasms in asymptomatic individuals. Eur J Gastroenterol Hepatol 2011;**23**(11):1036-41.

64. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;**105**(9):2017-25.

65. Sung JJ, Chan FK, Leung WK, et al. Screening for colorectal cancer in Chinese: comparison of fecal occult blood test, flexible sigmoidoscopy, and colonoscopy. Gastroenterology 2003;**124**(3):608-14.

66. Wong MCS, Ching JYL, Chan VCW, et al. Should prior FIT results be incorporated as an additional variable to estimate risk of colorectal neoplasia? A prospective study of 5,813 screening colonoscopies. Plos One 2014;**9**(12):e114332-e32.

67. Wu T, Kuo K, Wu Y, et al. Diagnostic accuracy of a single qualitative immunochemical fecal occult blood test coupled with physical measurements. Chin Med J (Engl) 2014;**127**(24):4164-70.

68. Bouvier V, Launoy G, Herbert C, et al. Colorectal cancer after a negative Haemoccult II(R) test and programme sensitivity after a first round of screening: The experience of the Department of Calvados (France). British Journal of Cancer 1999;**81**:305-09.

69. Castiglione G, Visioli C B, Ciatto S, et al. Sensitivity of latex agglutination faecal occult blood test in the Florence District population-based colorectal cancer screening programme. Br J Cancer

2007;**96**:1750-54.

70. Crotta S, Segnan N, Paganin S, et al. High Rate of Advanced Adenoma Detection in 4 Rounds of Colorectal Cancer Screening With the Fecal Immunochemical Test. Clin Gastroenterol Hepatol 2012;**10**:633-38.

71. Denters M J, Deutekom M, Bossuyt P M, et al. Lower risk of advanced neoplasia among patients with a previous negative result from a fecal test for colorectal cancer. Gastroenterology 2012;**142**:497-504.

72. Giai J, Exbrayat C, Boussat B, et al. Sensitivity of a colorectal cancer screening program based on a guaiac test: a population-based study. Clin Res Hepatol Gastroenterol 2014;**38**(1):106-11.

73. Itoh M, Takahashi K, Nishida H, et al. 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**:66-71.

74. Kronborg O, Fenger C, Sondergaard O, et al. Initial mass screening for colorectal cancer with fecal occult blood test. A prospective randomized study at Funen in Denmark. Scand J Gastroenterol 1987;**22**(6):677-86.

75. Launoy GD, Bertrand HJ, C. B. 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):493-96.

76. Levi Z, Birkenfeld S, Vilkin A, 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**:2415-24.

77. McNamara D, Leen R, Seng-Lee C, 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.

78. Nakama H, Kamijo N, Abdul Fattah A S, et al. Validity of immunological faecal occult blood screening for colorectal cancer: a follow up study. J Med Screen 1996;**3**:63-65.

79. Paimela H, Malila N, Palva T, et al. Early detection of colorectal cancer with faecal occult blood test screening. Br J Surg 2010;**97**:1567-71.

80. Parente F, Boemo C, Ardizzoia 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**:27-34.

81. Parra-Blanco A, Gimeno-Garcia AZ, Quintero E, et al. Diagnostic accuracy of immunochemical versus guaiac faecal occult blood tests for colorectal cancer screening. J Gastroenterol 2010;**45**(7):703-12.

82. Sieg A, Wirth A, Luthgens K, et al. Six years of screening for colorectal neoplasms with an immunological fecal occult hemoglobin test. Verdauungskrankheiten 2002;**20**:114-17.

83. Steele R J C, McClements P L, Libby G, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. Gut 2009;**58**:530-35.

84. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut 2013;**62**(3):409-15.

85. Doust JA, Pietrzak E, Sanders S, et al. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. J Clin Epidemiol 2005;**58**(5):444-49.

86. Buscemi N, Hartling L, Vandermeer B, et al. Single data extraction generated more errors than double data extraction in systematic reviews. J Clin Epidemiol 2006;**59**(7):697-703.

87. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. The New England journal of medicine 2012;**366**(8):687-96.

88. Naaktgeboren CA, de Groot JA, Rutjes AW, et al. Anticipating missing reference standard data when planning diagnostic accuracy studies. BMJ 2016;**352**:i402-i02.

89. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. Stat Med 2002;**21**(11):1525-37.

90. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med 2013;**158**(7):544-54.

91. Lee JK, Liles EG, Bent S, et al. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med 2014;**160**(3):171.

92. 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. British journal of cancer 2009;**100**(7):1103-10.

93. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;**135**(1):82-90.

CHAPTER 3

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

E.H. Schreuders*
E. Wieten*
S.A.V. Nieuwenburg
B.E. Hansen
I. Lansdorp-Vogelaar
E.J. Kuipers
M.J. Bruno
M.C.W. Spaander
*Both authors contributed equally to this manuscript.

Clinical Gastroenterology and Hepatology. 2016 Aug 24 (Epub ahead of print)

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).

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 χ^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.

RESULTS

Participant characteristics

The trial profile is summarized in Figure 3.1. 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

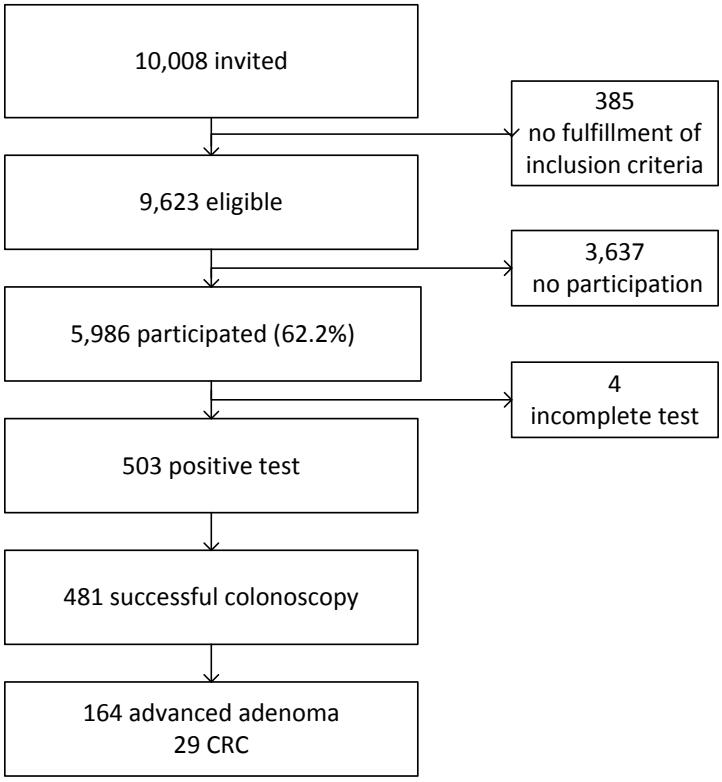
Test results per age category are given in Table 3.1. Screening participants had a median age of 61 years and 48% of the participants was male. Attendance rate and positivity rate increased significantly with age ($P<.001$). Detection rates of AN and CRC increased significantly with age as continuous 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 3.1 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.

Figure 3.1 Trial profile.



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 3.2). 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

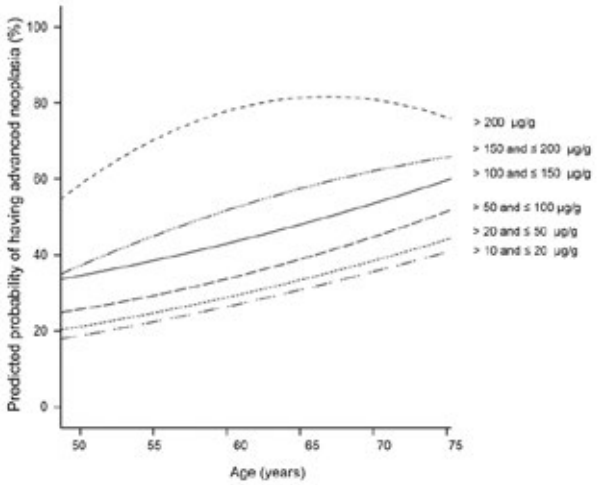
between the observed values and the predicted values (Goodness-of-Fit; *P*=.276). The predicted probability of having AN for participants with a positive FIT per age, corrected for gender and fecal Hb concentration, was depicted in Figure 3.2 for different fecal Hb concentration subgroups.

Table 3.2 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

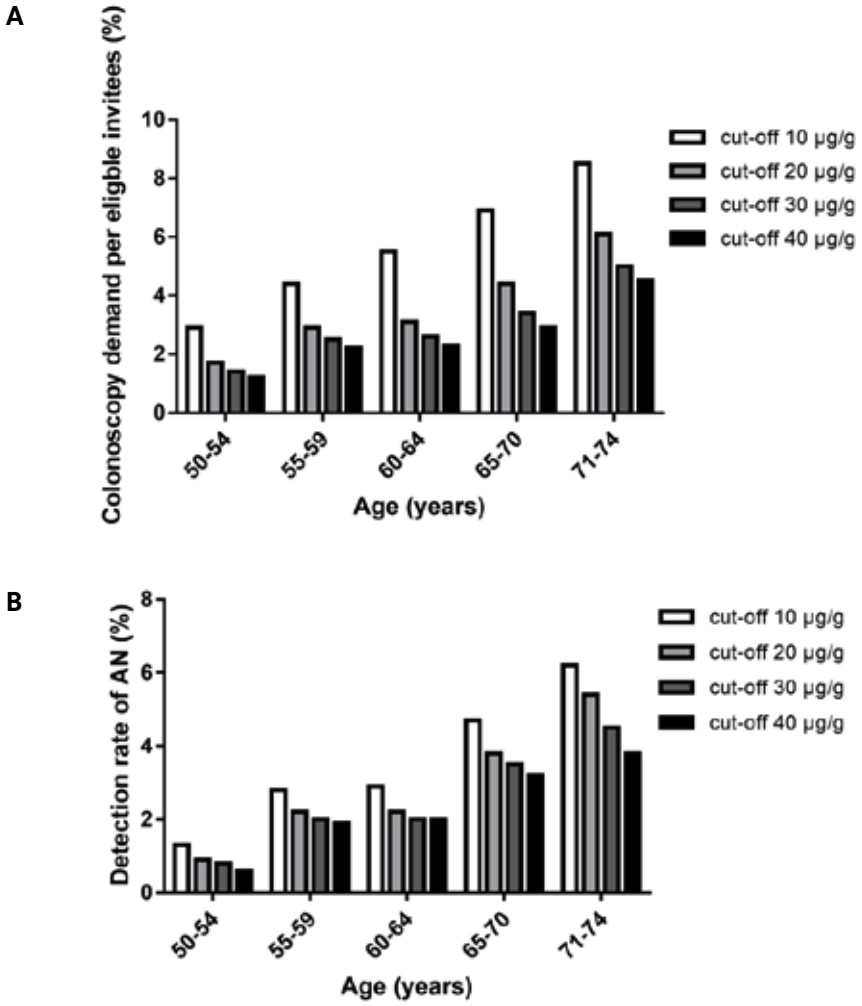
FIT: fecal immunochemical test; CRC: colorectal cancer.

Figure 3.2 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.

Figure 3.3 A. Colonoscopy demand and B. detection rate of advanced neoplasia (AN) per age category and cut-off concentration in µg Hb/gram feces.



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 3.3a and Figure 3.3b. 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 3.3). 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 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.

Table 3.3 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 (95% CI)
	Age range, cut-off **	N (%*)	AN (%*)	CRC (%*)	
	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
** (µg Hb/g feces)
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.

Screening strategies with age-specific cut-off concentrations were also assessed on these outcomes (Table 3.3). 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.

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 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 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, et al. Global cancer statistics. CA: a cancer journal for clinicians 2011;**61**(2):69-90.
2. Hewitson P, Glasziou P, Irwig L, et al. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database of Systematic Reviews 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, 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, 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, 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, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. British journal of cancer 2009;**100**(7):1103-10.
8. Brenner H, Hoffmeister M, Stegmaier C, et al. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 2007;**56**(11):1585-9.
9. Ferlitsch M, Reinhart K, Pramhas S, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA 2011;**306**(12):1352-8.
10. van Rossum LG, van Rijn AF, Laheij RJ, et al. Cutoff value determines the performance of a semi-quantitative immunochemical faecal occult blood test in a colorectal cancer screening programme. British journal of cancer 2009;**101**(8):1274-81.
11. 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 2012;**50**(5):935-40.
12. Fraser CG, Rubeca T, Rapi S, et al. 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, 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, 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, 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, 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: a cancer journal for clinicians 2008;**58**(3):130-60.

17. Auge JM, Pellise M, Escudero JM, et al. Risk Stratification for Advanced Colorectal Neoplasia According to Fecal Hemoglobin Concentration in a Colorectal Cancer Screening Program. Gastroenterology 2014.

18. Wilschut JA, Hol L, Dekker E, 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, et al. Gender differences in faecal haemoglobin concentration. J Med Screen 2016;**23**(1):54.

21. 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. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver 2016;**48**(5):542-51.

22. Kapidzic A, van der Meulen MP, Hol L, et al. Gender Differences in Fecal Immunochemical Test Performance for Early Detection of Colorectal Neoplasia. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2015;**13**(8):1464-71 e4.

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, et al. Multitarget stool DNA testing for colorectal-cancer screening. The New England journal of medicine 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, et al. Histopathological assessment of bleeding from polyps of the colon and rectum. J Clin Pathol 1989;**42**(4):378-82.

CHAPTER 4

Comparison of multiple rounds one versus two-sample fecal immunochemical test-based colorectal cancer screening

E.H. Schreuders
E.J. Grobbee
A. Kapidzic
A.H.C. van Roon
A.J. van Vuuren
I. Lansdorp-Vogelaar
K. Izelaar
M.J. Bruno
E.J. Kuipers
M.C.W. Spaander

Submitted

ABSTRACT

Background

Many countries implement fecal immunochemical test (FIT) based colorectal cancer (CRC) screening. Because of suboptimal sensitivity of a single FIT, successive rounds are required to achieve an optimal preventive effect. It is unknown whether 2-sample FIT (2-FIT) increases program sensitivity over multiple screening rounds. Therefore we compared CRC screening by repeated 2-FIT to one sample FIT (1-FIT) on participation, positive predictive value (PPV), diagnostic yield and interval CRC.

Methods

A random selection of 13,205 average-risk individuals in The Netherlands, aged 50–74 years, were between 2006 and 2015 invited for four rounds of biennial 1-FIT or 2-FIT screening. Per round, persons received one or two identical FITs to sample on one versus two consecutive bowel movements. Screenees with a hemoglobin (Hb) concentration of ≥10 μg Hb/g feces in at least one test were referred for colonoscopy. Persons with a positive FIT in previous rounds were not re-invited for FIT screening.

Results

In total, 9,787 1-FIT and 3,131 2-FIT invitees were eligible at least once during the screening period. In the 1-FIT screening cohort, 75% participated at least once versus 73% in the 2-FIT cohort ($p=0.01$). Of persons attending at least once, the cumulative positivity rate was 19% for 1-FIT and 29% for 2-FIT screening ($p<0.001$). Cumulative PPV for advanced neoplasia was significantly lower for 2-FIT screening, 24% compared to 33% for 1-FIT screening ($p<0.001$). The cumulative yield for advanced neoplasia after 4 rounds was 4.4% for 1-FIT and 4.7% for 2-FIT screening ($p=0.46$). The interval CRC rates for 1-FIT and 2-FIT screening were 0.08% and 0.06% respectively.

Conclusions

Repeated 1-FIT and 2-FIT screening result in the same yield of advanced neoplasia. The lower PPV for 2-FIT screening results in a higher colonoscopy demand. Therefore, the use of 1-FIT screening is recommended for FIT-based CRC screening programs.

BACKGROUND

Colorectal cancer (CRC) ranks third among all cancers worldwide, affecting approximately 1.36 million patients each year.[1] The high burden of disease and long preclinical stage make CRC suitable for population screening .[2 3] Long-term randomized prospective trials using guaiac fecal occult blood testing (gFOBT) have proven to decrease CRC-related mortality.[4 5] Currently many countries are implementing fecal immunochemical test (FIT) based CRC screening.[6] These programs are routinely based on one sample from one bowel movement, whereas gFOBT uses 6 samples from three consecutive bowel movements. Advanced neoplasia can bleed intermittently and may thus be missed with single stool sampling. Screening by means of 2-sample FIT (2-FIT) can reduce the risk of missing advanced lesions, and increase test sensitivity.[7]

We and others demonstrated that, in a single screening round, 2-FIT screening at a low cut-off detects significantly more advanced neoplasia than 1-sample FIT (1-FIT).[8-10] However, FIT screening requires successive rounds for an optimal effect. Results after repeated 1-FIT screening have been published.[11-13] There are however currently no data on multiple rounds of 2- FIT screening, nor comparisons of both strategies.

Interval cancers, i.e. cancers occurring after a negative screening test and before the next screening test is due, are worldwide considered an important indicator of the quality and effectiveness of CRC screening.[14] When evaluating multiple rounds of screening, assessment of interval cancer rates is of paramount importance. Although large number of screening studies were performed over the last two decades, few studies reported on interval CRC rates.[15-17]

We therefore conducted a population-based CRC screening trial to determine participation, positive predictive value (PPV) diagnostic yield and interval cancer rate of 2-FIT versus 1-FIT screening over four successive rounds.

METHODS

Population and design

The screening cohort have been described elsewhere.[8 11 18] In short, 13,205 average-risk individuals were invited for CRC screening by means of either 1-FIT or 2-FIT screening. Demographic data of all individuals between 50 and 74 years living in the Rotterdam-Rijnmond region in the Netherlands were obtained from

municipal population registers. Random samples of the selected population were taken based on different postal codes by a computer-generated algorithm (Tenalea, Amsterdam, The Netherlands). Allocation to 1-FIT or 2-FIT screening occurred prior to invitation. As there was no CRC screening program at the time of the trial, the cohort was screening-naive when first approached. The screening cohort was invited by postal mail and invitations were coordinated by and sent from the Regional Organization for Population Screening in the South-West of the Netherlands.

Per round, persons were excluded if they had a history of CRC or inflammatory bowel disease, had undergone colon imaging ≤ 3 years, had an estimated life expectancy <5 years, were unable to give informed consent, had been positive in previous rounds, had become older than 74 years, had moved out of the region, or when they had died.

1-FIT screening cohort

The 1-FIT screening cohort (n=10,008) was invited from 2006 to 2014 for four rounds of biennial 1-FIT screening. After the first round, this cohort was randomized over three groups to repeat screening at different intervals in the second round (i.e., one, two, and three years, respectively).[17] No significant difference in participation, detection rate or positive predictive value (PPV) was found between groups.[17] Therefore, a two year interval was applied to all groups in the third and fourth screening round. With each screening round, one FIT (OC-sensor Micro, Eiken Chemical, Tokyo, Japan) was sent to collect a single sample of one bowel movement. At the fourth round, this cohort was randomized to receive either an OC-sensor or an FOB-Gold (Sentinel, Italy) at the same fecal Hb cut-off concentration (trial registration no. NTR5385).[19] The OC-Sensor and FOB-Gold perform equally over the relevant concentration range.[20]

2-FIT screening cohort

The 2-FIT screening cohort (n=3,197) was invited from 2008 to 2015 for four rounds of biennial 2-FIT screening (trial registration no. NTR5740). Each round, participants received two identical FITs (OC-sensor Micro, Eiken Chemical, Tokyo, Japan) to sample from two consecutive bowel movements.

FIT analysis and follow-up

For both cohorts, the test result was considered positive if the hemoglobin concentration in the FIT sample was ≥ 10 µg Hb / g feces in at least one FIT. The Fecal Immunochemical TesTs for Hemoglobin Evaluation Reporting (FITTER) check list was used to adequately report handling of the FIT (Supplementary file 4.1).[21] Screenees with a positive FIT were referred to colonoscopy. Experienced

endoscopists performed all colonoscopies. The maximum reach of the endoscope, adequacy of bowel preparation, and the characteristics and location of any polyps were recorded. Location was considered right-sided when proximal to the splenic flexure. Colonoscopy results were classified according to the most advanced lesion found. All removed polyps were evaluated by experienced gastrointestinal pathologists. Advanced neoplasia included CRC and advanced adenomas. An advanced adenoma was defined as an adenoma >10mm, or >25% villous component and/or high-grade dysplasia. Patients with a positive colonoscopy entered a surveillance program, whereas persons with a negative colonoscopy were excluded from FIT screening for 10 years.[22]

In the Netherlands all cancers are registered by the Dutch Comprehensive Cancer Centre (www.iknl.nl). For our study all recruited participants were linked with the Dutch Comprehensive Cancer Centre.

Definitions

The participation rate was defined as the number of eligible participants relative to all eligible invitees (eligible defined as all invitees minus the excluded invitees). The positivity rate was defined as the proportion of positive tests in eligible participants with an analyzable test-result. Detection rate was defined as persons with advanced neoplasia or CRC relative to all participants. Diagnostic yield was defined as those with advanced neoplasia or CRC relative to all invitees. The PPV compromised all diagnosed with advanced neoplasia or CRC proportionally to positive screenees who underwent colonoscopy. The number needed to scope was calculated as the number of colonoscopies needed to find one screenee with advanced neoplasia. Number needed to screen describes the number of complete FITs needed to find one case with advanced neoplasia.

Screen-detected CRC was defined as CRC diagnosed at colonoscopy performed after a positive FIT. Interval cancers were defined as CRC detected after a negative FIT and before the next invitation.[14] A CRC after a positive FIT and subsequent colonoscopy but before the next colonoscopy is due, was considered a post-colonoscopy interval CRC (within a FIT screening program).[14]

Statistical analysis

Differences in proportions were calculated using a Chi-squared test and expressed with corresponding 95% confidence intervals (CI). Differences in means were analyzed using Student's T test. To test differences in participation between rounds, a generalized estimating equation was used to account for clustering at the level of the invitee. The cumulative screening results after four rounds of 1-FIT and 2-FIT screening were calculated. Cumulative participation

was calculated as the total number of persons participating at least once, relative to all persons who had been eligible at least once over four screening rounds. This approach most closely mimics a population-based screening program. Cumulative yield reflected the total number of persons with advanced neoplasia or CRC in all persons who had been eligible at least once. The program sensitivity, expressed as a percentage, was calculated by dividing the number of screenees with screen-detected CRC, by the total number of screenees diagnosed with both interval CRC and screen-detected CRC. For all tests, a two-sided-p-value of <0.05 was considered significant. SPSS package version 22.0 was used for the calculations.

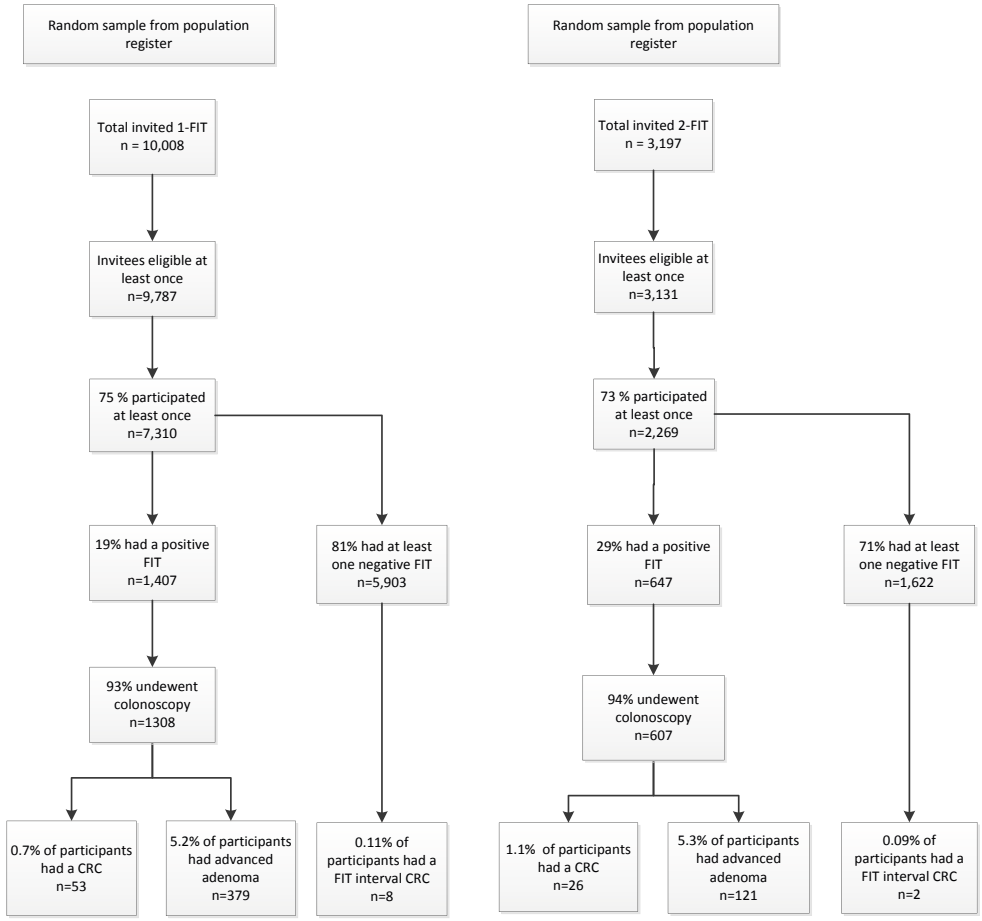
The Dutch National Health Council and the Institutional Review Board of the Erasmus MC University Medical Centre approved the study. All participating screenees gave written informed consent.

RESULTS

Four rounds of 1-FIT and 2-FIT screening

Over four rounds, a total of 10,008 persons were invited for 1-FIT screening and 3,197 persons were invited for 2-FIT screening (Figure 4.1). The results per round are shown in Table 4.1, showing high and stable participation rates in both screening cohorts. In the first round, 8.4% of test results were positive in the 1-FIT cohort and 12.7% in the 2-FIT cohort ($p<0.001$). Positivity rates of both cohorts decreased over subsequent rounds (Table 4.1). The positivity rate of the 2-FIT cohort rose between the third and fourth round from 9.6% to 10.8%, although this was not significantly different ($p=0.367$). In all rounds, the PPV of FIT for advanced neoplasia was significantly lower for 2-FIT screening compared to 1-FIT screening (Table 4.1). Differences in diagnostic yield among eligible invitees between 1-FIT and 2-FIT screening are shown in Figure 4.2. In the first round, the number needed to screen to detect one subject with advanced neoplasia was lower for the 2-FIT cohort; 40 versus 50 for the 1-FIT cohort. In subsequent rounds, this was reversed (Table 4.1).

Figure 4.1 Flow chart



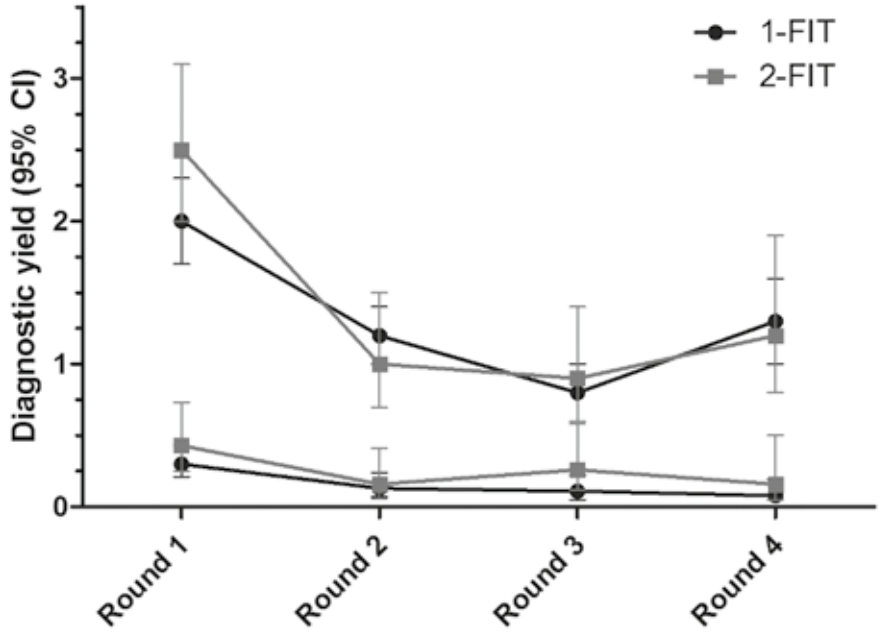
CRC: colorectal cancer, FIT; fecal immunochemical test

Table 4.1 Four rounds 1-FIT versus 2-FIT screening

	Round 1		Round 2		Round 3		Round 4	
	1-sample	2-sample	1-sample	2-sample	1-sample	2-sample	1-sample	2-sample
Eligible invitees	9,623	3,057	8,158	2,579	7,326	2,287	6,452	1,845
Participated, n (%)	5,986 (62.2)	1,875 (61.3)	5,200 (63.7) [b]	1,582 (61.3) [a]	4,998 (68.2) [b]	1,474 (64.4) [a b]	4,385 (64.7) [b]	1,171 (63.5) [a]
Positive test, n (%)	503 (8.4)	239 (12.7) [a]	299 (5.8) [b]	132 (8.3) [a b]	275 (5.5) [b]	143 (9.6) [a]	322 (7.3) [b]	126 (10.8) [a]
Colonoscopy compliance	481 (96)	226 (95)	289 (97)	128 (97)	255 (93)	137 (96)	282 (88)	116 (92) [a]
Detection rate, n (%)								
AN	193 (3.2)	77 (4.1) [a]	97 (1.9) [b]	27 (1.7) [a b]	59 (1.2) [b]	20 (1.4) [a]	83 (1.9) [b]	23 (2.0) [a]
CRC	29 (0.5)	13 (0.7)	11 (0.2) [b]	4 (0.3)	8 (0.2)	6 (0.4)	5 (0.1)	3 (0.3)
PPV, % (95% CI)								
AN	40 (36 – 45)	34 (28 – 41)	34 (28 – 39)	21 (15 – 29)	23 (18 – 29)	15 (10 – 22) [a]	29 (24 – 45)	20 (12 – 28) [a]
CRC	6.0 (4.2 – 8.5)	5.8 (3.4 – 9.7)	3.8 (2.1 – 6.7)	3.1 (1.2 – 8.0)	3.1 (1.6 – 6.1)	4.4 (2.0 – 9.4)	1.8 (0.7 – 4.2)	2.6 (0.8 – 7.7)
Number needed to screen to detect one case with AN	50	40	84	96	124	114	78	80
Number needed to scope to detect one case with AN	2.5	2.9	3.0	4.7	4.3	6.9	3.4	5.0

[a] $P<0.05$ compared with 1-FIT screening in the same round
[b] $P<0.05$ compared with the previous round
AN; advanced neoplasia, CRC; colorectal cancer, PPV; positive predictive value, CI; confidence interval.

Figure 4.2 Diagnostic yield advanced neoplasia (upper lines) and CRC (lower lines) among eligible invitees



Cumulative results after four rounds of 1-FIT and 2-FIT screening

Over four rounds, 9,878 screenees from the 1-FIT cohort and 3,131 screenees from the 2-FIT cohort were eligible at least once. In total, 75% of these eligible persons in the 1-FIT screening cohort participated at least once, versus 73% in the 2-FIT cohort (p=0.01). Among those attending at least once, the cumulative positivity rate of 1-FIT and 2-FIT screening was 19% (95%CI; 18 – 20) and 29% (95%CI; 27 – 30) respectively (P<0.001). The cumulative detection rate of advanced neoplasia among participants was 6.0% (95%CI; 5.5 – 6.6) for 1-FIT and 6.6% (95%CI; 5.6 – 7.7) for 2-FIT screening (p=0.32). The cumulative PPV for advanced neoplasia was significantly lower for 2-FIT screening, 24% compared to 33% for 1-FIT screening (p<0.001). The cumulative yield of CRC among invitees was 0.7% (95%CI; 0.6 – 1.0) for 1-FIT and 1.2% (95%CI; 0.8 – 1.7) for 2-FIT screening (p=0.06).

Screen detected and interval carcinomas

In the 1-FIT screening cohort, a total of 89 (1.2%) CRCs were detected among participants within the study period. Of these, 53 were screen-detected at colonoscopy after a positive FIT. Among screenees that underwent colonoscopy because of a positive 1-FIT, seven post-colonoscopy interval CRCs were found (0.5%). Classification of other detected CRCs are shown in Table 4.2. Eight

persons (0.08% of screenees that were eligible at least once during the study period) were diagnosed with an interval carcinoma after a negative FIT.

In the 2-FIT screening cohort, a total of 35 CRCs (1.5% of participants) were detected of which 26 were screen detected. Two FIT interval CRCs were found compromising 0.06% of screenees that were eligible at least once during the study period (p=1.00 compared to 1-FIT). Among screenees that underwent colonoscopy because of a positive 2-FIT, one post-colonoscopy interval CRC was found (0.3%).

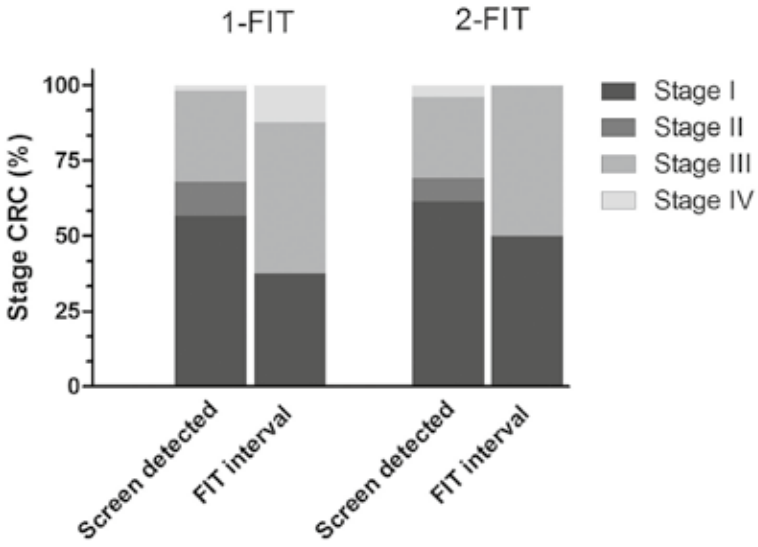
Among screenees participating at least once, screen-detected CRC rate was 0.7% for 1-FIT and 1.1% for 2-FIT screening (p=0.06). Screen-detected CRCs were more often left-sided in both cohorts, 74% of 1-FIT and 85% of 2-FIT (P=0.27). FIT interval CRC rate among screenees participating at least once was 0.11% for 1-FIT and 0.09% for 2-FIT screening (p=1.00). FIT interval CRCs were equally left and right sided in both cohorts (both 50%). Among screenees participating at least once in the 1-FIT cohort, program sensitivity for CRC was 60% (95%CI 49 – 69). The program sensitivity for CRC in the 2-FIT participants was 74% (95%CI 58 – 86).

Table 4.2 Characteristics of CRCs detected among invitees eligible at least once

	1-sample N=9,787	**2-sample N=3,131**
	n (%)	n (%)
Screen detected	53 (0.54)	26 (0.83)
FIT interval carcinoma	8 (0.08)	2 (0.06)
Positive FIT but refused colonoscopy	1 (0.01)	1 (0.03)
Post-colonoscopy interval carcinoma	7 (0.07)	1 (0.03)
Surveillance detected	3 (0.03)	2 (0.06)
Not eligible*	17 (0.17)	3 (0.09)
Total	89	35

* Includes CRC detected before first invitation, and in screenees older than screening age range.

Figure 4.3 Distribution of CRC by stage for 1-FIT and 2-FIT screening



DISCUSSION

Currently many countries are implementing FIT based CRC screening programs. To optimize FIT based CRC screening programs, to meet financial and colonoscopy resources, different strategies can be followed. The suboptimal sensitivity of FIT for detection of advanced neoplasia asks for repeated screening at 2-year interval. This is the first study to compare 1-FIT versus 2-FIT screening over multiple rounds. Participation in repeated rounds of FIT screening was high for both 1-FIT and 2-FIT screening. After four rounds, 1-FIT and 2-FIT screening result in the same yield of advanced neoplasia. The FIT interval cancer rates were 0.08% for 1-FIT and 0.06% for 2-FIT screening.

In our population based screening cohorts, stable and high participation rates over multiple rounds were found. Cumulative participation for 1-FIT and 2-FIT screening was 75% and 73% respectively (p=0.01). An American study, using 1-FIT, reported a total participation of 63.8% over 4 rounds.[13]

The FIT positivity and detection of advanced neoplasia were highest in the first round. This pattern is consistent with detecting more prevalent cases of advanced neoplasia in the initial round of screening-naïve persons and more incident cases in subsequent rounds. Similar trends have been reported in previous FIT and gFOBT screening trials.[11 13 23 24] An increase in positivity rate in the fourth round

was also seen in other FIT screening trials.[12 13] Possible explanations for this rise are increasing age, improvement in FIT buffer, and the fact that for this analyses a closed cohort (i.e., which was not replenished with screening naïve persons) was used.

Effective population based FIT screening requires follow-up colonoscopy after a positive FIT result. In total, 93% to 94% of positive screenees received colonoscopy after a positive result. This is above the internationally accepted quality indicator of more than 90% adherence to colonoscopy after a positive FIT.[25] While our results show high adherence rates, low colonoscopy follow-up rates have also been reported in literature.[26 27] Low colonoscopy follow-up rates should be assessed to increase CRC screening effectiveness since positive screenees are at high risk of having advanced neoplasia.[11]

Although numbers are small, the relative high percentage of stage one CRCs aligns with the expectation of ongoing detection of newly developing cancers during aging (Figure 4.3). Nevertheless, interval cancers after a negative FIT also occurred, yet to a lesser extent with 2-FIT versus 1- FIT screening, although not statistically significant.. Theoretically, 2-FIT screening may detect tumors that bleed intermittently more easily because fecal samples from bowel movements of different days are used. Interval cancers after a negative FIT remain an on-going concern and also suggest different biology compared to screen detected CRC. The so-called serrated pathway has been proposed as an alternative to the adenoma-carcinoma sequence, and is thought to contribute to the risk of interval CRC, especially in the proximal colon.[28 29]

In the first round, detection rates for advanced neoplasia and CRC, as well as number needed to screen and scope to detect one subject with advanced neoplasia, were higher for 2-FIT than 1-FIT screening. However, after four consecutive rounds of FIT screening, the diagnostic yield among invitees was similar for both cohorts, with 4.4% for 1-FIT, and 4.7% for 2-FIT ($p=0.46$). Switching from a first round with 2-FIT to a second round with 1-FIT screening has been investigated after the second round.[18] The additional value of a second test in the second round seemed limited, since more colonoscopies were required to detect additional advanced neoplasia. Modelling studies are necessary to evaluate cost-effectiveness of 1-FIT versus 2-FIT screening. This can aid in defining the optimal FIT screening strategy by simulating different cut-offs and screening intervals for both strategies.

Potential limitations of our study include the relatively small sample size of the 2-FIT cohort. In addition, persons were not randomized between 1-FIT and 2-FIT

screening. The fact that both study populations were randomly drawn from the same municipal registries in the same region, within the same time period indicates that both populations can be compared head to head.

Worldwide, different screening modalities and strategies exist, often tailored to available financial resources and colonoscopy capacity.[6] Countries offering a FIT-based screening program vary in screening interval, cut-off concentrations and eligibility criteria in terms of age to start and stop screening.[6] However, the optimal number of FITs to be used per screening round has not yet been investigated. We showed that after four rounds, 1-FIT versus 2-FIT screening yielded no difference in detection of advanced neoplasia and CRC. In this regard, the overall diagnostic yield of advanced neoplasia over time is a strong indicator to determine the impact of a screening program, especially to compare different screening tests or strategies. Notably, 2-FIT screening requires a substantial higher colonoscopy capacity due to a lower PPV. Given the limited colonoscopy capacity in most countries, the use of 1-FIT screening is recommended.

REFERENCES

1. GLOBOCAN. Estimated cancer incidence, mortality and prevalence worldwide in 2012 http://globocan.iarc.fr/Default.aspx, 2012.
2. Kuntz KM, Lansdorp-Vogelaar I, Rutter CM, et al. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. Med Decis Making 2011;**31**(4):530-9.
3. Brenner H, Hoffmeister M, Stegmaier C, et al. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 2007;**56**(11):1585-9.
4. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;**369**(12):1106-14.
5. Hewitson P, Glasziou P, Irwig L, et al. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev 2007(1):CD001216.
6. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.
7. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;**105**(9):2017-25.
8. van Roon AH, Wilschut JA, Hol L, et al. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. Clin Gastroenterol Hepatol 2011;**9**(4):333-9.
9. Raginel T, Puvinel J, Ferrand O, et al. A population-based comparison of immunochemical fecal occult blood tests for colorectal cancer screening. Gastroenterology 2013;**144**(5):918-25.
10. Guittet L, Bouvier V, Mariotte N, et al. Performance of immunochemical faecal occult blood test in colorectal cancer screening in average-risk population according to positivity threshold and number of samples. Int J Cancer 2009;**125**(5):1127-33.
11. Kapidzic A, Grobbee EJ, Hol L, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. Am J Gastroenterol 2014;**109**(8):1257-64.
12. Crotta S, Segnan N, Paganin S, et al. 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.
13. Jensen CD, Corley DA, Quinn VP, 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.
14. 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.
15. Steele RJ, McClements P, Watling C, 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.
16. Denters MJ, Deutekom M, Bossuyt PM, et al. 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.
17. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based

colorectal cancer screening. Gut 2013;**62**(3):409-15.

18. Kapidzic A, van Roon AH, van Leerdam ME, et al. Attendance and diagnostic yield of repeated two-sample faecal immunochemical test screening for colorectal cancer. Gut 2015.

19. Grobbee EJ, van der Vlugt M, van Vuuren A, et al. Comparison of OC-Sensor and FOB-Gold in Population-Based Colorectal Cancer Screening Based on FIT. Gastroenterology 2015;**148**(4):S-160.

20. Grobbee EJ, van der Vlugt M, van Vuuren A, et al. Comparison of OC-Sensor and FOB-Gold in Population-Based Colorectal Cancer Screening Based on FIT. Gastroenterology;**148**(4):S-160.

21. Fraser CG, Allison JE, Young GP, et al. 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.

22. Levin B, Lieberman DA, McFarland B, 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.

23. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;**348**(9040):1472-7.

24. Steele RJ, McClements PL, Libby G, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. Gut 2009;**58**(4):530-5.

25. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, 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.

26. Plumb AA, Ghanouni A, Rainbow S, et al. Patient factors associated with non-attendance at colonoscopy after a positive screening faecal occult blood test. J Med Screen 2016.

27. Ferrat E, Le Breton J, Veerabudun K, et al. Colorectal cancer screening: factors associated with colonoscopy after a positive faecal occult blood test. Br J Cancer 2013;**109**(6):1437-44.

28. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010;**138**(6):2088-100.

29. Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 2012;**107**(9):1315-29; quiz 14, 30.

SUPPLEMENTARY FILES

Supplementary file 4.1 The FITTER check-list for the reporting of studies using Fecal Immunochemical Tests for hemoglobin (FIT) [21]

Specimen Collection and Handling
In the first 3 rounds subjects received the OC-sensor (Eiken, Japan), and in the last round only 1-FIT cohort subjects received either the OC-sensor or the FOB-Gold (Sentinel, Italy). Mean mass of feces collected by applicator stick:

OC-sensor: 10 mg feces, collected in 2 mL buffer.
FOB-Gold: 10 mg feces, collected in 1.7 mL buffer.

Participants were asked to sample feces according to instructions and post the feces sample within 24 hours after collection or keeping in the refrigerator until mailing. Participants signed an informed consent and were asked to write the date of sample collection on the device label before returning by post. They returned the FIT(s) and at ambient temperature by freepost to the laboratory of the Erasmus Medical Centre.

Analysis
The tests were stored at -20° until analysis, for at most 14 days. The OC sensor FITs were analyzed on the OC-sensor μ systems (Eiken, Japan), the FOB-gold FITs were analyzed on a Sentinel Sentifit 270 system (Sentinel, Italy). All FIT tests were allowed to warm to room temperature before analysis and analyzed once. The analytical working range for the OC sensor μ was 1-200 μg Hb/g feces, and 1-170 μg Hb/g feces for the Sentifit 270. Samples with fecal Hb concentrations above the upper analytical working limits were not diluted or re-analyzed.

Quality Management
Analytic runs were accepted only if the calibration and controls were in the margins. The analyses were performed by 4 trained laboratory analysts. After the test was analyzed, the laboratory personnel entered the test result into the database.

Result Handling
Units used, with conversion to μg Hb/g feces if ng Hb/mL was not used:

OC-sensor: 50 ng/mL Hb χ10 μg Hb/g feces.
FOB-Gold: 59 ng/mL Hb χ10 μg Hb/g feces.

The test result was considered positive if the Hb concentration in the FIT sample was ≥ 10 μ g Hb / g feces in at least one FIT.

CHAPTER 5

Fecal haemoglobin concentrations predict future advanced colorectal neoplasia in long-term population-based FIT-screening

E.J. Grobbee
E.H. Schreuders
B.E. Hansen
M.J. Bruno
I. Lansdorp-Vogelaar
M.C.W. Spaander
E.J. Kuipers

Submitted

ABSTRACT

Background and aims

Colorectal cancer (CRC) screening using quantitative fecal immunochemical tests (FITs) is rapidly gaining ground worldwide. FITs are invariably used in a dichotomous manner using pre-specified cut-offs. To optimize FIT-based screening programs, we explored if fecal hemoglobin concentrations ($_f$Hb) of participants with a FIT below the cut-off (FIT_{bco}) could be used to predict future colorectal advanced neoplasia (AN) risk.

Methods

Average-risk subjects aged 50-74 years, were offered four rounds of population-based FIT screening (cut-off 10 µg Hb/g feces). All subjects with a FIT_{bco} at first participation (baseline) were included. Hazard ratios (HRs) for AN were determined using Cox proportional hazard regression analyses. Logistic regression techniques were used to calculate risks of AN after consecutive FIT_{bco} results.

Results

Out of 13,566 invitees, 9,561 (70%) participated at least once and 7,663 (92%) had FIT_{bco} at baseline. Median follow up was 4.7 years (IQR 2.0-6.1). After eight years of follow-up, a higher cumulative incidence of AN was found in screenees with baseline $_f$Hb between 8 to 10 µg Hb/g compared to those with $_f$Hb of 0 µg Hb/g (5 vs. 33%; $p<0.001$). The multivariate HRs increased from 1.2 to 8.2 for $_f$Hb concentrations between >0 and 2 µg Hb/g and ≥8 to 10 µg Hb/g ($p<0.001$). A 14-fold increased risk was found after two consecutive $FITs_{bco}$ with twice $_f$Hb of 8 µg Hb/g versus twice 0 µg Hb/g ($p<0.001$).

Conclusion

Among screenees with a FIT_{bco}, baseline and consecutive $_f$Hb are independent predictors for incident AN. These findings provide tools for personalized strategies in population-based CRC screening. This may decrease unnecessary screenee burden and could optimize use of resources.

INTRODUCTION

Colorectal cancer (CRC) is among the most common causes of cancer-related mortality.[1] Population-based CRC screening can significantly reduce disease burden. Fecal occult blood tests are widely accepted for this purpose.[2 3] A higher fecal hemoglobin ($_f$Hb) concentration is associated with a higher risk of advanced neoplasia (AN).[4-7] Many screening programs worldwide use fecal immunochemical tests (FIT), that can be either qualitative (i.e. providing a positive or negative test result) or quantitative (i.e. quantifying $_f$Hb concentrations in feces).[2 8] Although quantitative FITs provide exact $_f$Hb concentrations in µg Hb/g feces, current screening programs routinely use $_f$Hb in a dichotomized fashion. As such, they are invariably used as qualitative tests. A test is considered positive above a fixed threshold that is the same for all screenees in all rounds of screening. Those with a positive test are recommended to undergo colonoscopy. Individuals with a negative test are offered a renewed FIT after a predefined screening interval without taking into account previous $_f$Hb concentrations. Most FIT screening programs rely on annual or biennial screening, requiring participants to repeat the test multiple times over the course of years.

To increase screening efficiency and impact of FIT screening programs, it is relevant to explore if screenees with a negative FIT, that is a $_f$Hb concentration below the pre-defined cut-off level (FIT below cut-off; FIT_{bco}), can be categorized according to their actual $_f$Hb concentration into different risk groups for later development of AN. Such tailored screening would allow for targeted variation of screening intervals, and decrease screening and colonoscopy demand or optimize its use. Currently, many countries with CRC screening programs struggle to match colonoscopy demands with limited resources.[9-11] Over the course of multiple screening rounds, $_f$Hb concentration could then be of guidance in identifying those at low and high risk of AN, and thus form the basis of individualized screening strategies. Such information is of key importance for national population-based CRC screening policies. Previous studies have mainly focused on $_f$Hb concentrations of FIT-positive screenees, or have only assessed first round $_f$Hb concentrations.[5 12 13] At present, no literature is available on trends in individual $_f$Hb concentrations of FIT_{bco} screenees over consecutive screening rounds. Furthermore, it is not known whether these previous $_f$Hb concentration(s) can be used as a predictor for the future detection of AN. Therefore, we aimed to investigate trends in $_f$Hb of FIT_{bco} results at first participation, and in subsequent rounds as a predictor for future incidence of AN.

METHODS

Study design and participants

Details about this study cohort and the design have been described before.[14] In short, individuals living in the southwest of the Netherlands were approached for four rounds of FIT-screening. Demographic data of all individuals between 50 and 74 years living in this region were obtained from municipal population registers. Random samples were taken based on different postal codes. Exclusion criteria included a history of inflammatory bowel disease or CRC, a full colonic examination in the past two years, an estimated life expectancy of <5 years, and inability to give informed consent. In case of a positive FIT result, subjects were sent for colonoscopy and not re-invited for subsequent FIT rounds. Subjects were not invited when they had become older than 74 years, or when they had moved out of the region. The cohort was supplemented with new screening-naïve subjects in round 3 and 4 to best mimic a continuous, population-based screening cohort. Recruitment took place between November 2006 and December 2014. For this analysis all subjects participating in at least one screening round and with a negative FIT result, defined as a FIT result below the cut-off of 10 µg Hb/ g feces (FIT_{bco}) at their first participation (i.e. baseline), were included.

FIT screening and colonoscopy

Each screening round, eligible invitees received one FIT per mail. Invitees were instructed to collect one sample of one bowel movement. In the first 3 rounds subjects received the OC-sensor (Eiken, Japan), and in the last round subjects were randomized to receive either the OC-sensor or the FOB-Gold (Sentinel, Italy). The OC-Sensor and FOB-Gold perform equally over the relevant concentration range.[15] Participants were asked to sample feces according to instructions and post the sample together with the consent form within 24 hours while storing it in the refrigerator until mailing. The OC sensor FITs were analyzed on the OC-sensor µ system (Eiken, Japan), the FOB-gold FITs were analyzed on a Sentifit 270 system (Sentinel, Italy). All FIT tests were analyzed once at room temperature. The analytical working ranges for the OC sensor µ and Sentifit 270 were respectively 1-200 feces and 1-170 µg Hb/g feces. Samples with $_{f}$Hb concentrations above the upper analytical working limits were not diluted or re-analyzed. The test result of ≥10 µg Hb/g feces was considered positive. Subjects with a positive FIT were scheduled for colonoscopy within 4 weeks. In case colonoscopy was incomplete, a computed tomographic colonography was performed. Experienced board-certified gastroenterologists performed all endoscopies. The maximum reach of the endoscope, adequacy of bowel preparation, and characteristics and location of any polyps were recorded. All

polyps were removed and evaluated by dedicated gastrointestinal pathologists. Patients with a positive colonoscopy entered a surveillance program according to guidelines of the Dutch Society of Gastroenterology, whereas subjects with a negative colonoscopy were considered not to require FIT screening for 10 years. AN was defined as an adenoma of ≥ 10 mm, an adenoma with at least 25% villous histology and/or high-grade dysplasia, or CRC.

Follow-up and interval carcinomas

Except for individuals who had moved out of the Netherlands, all recruited participants were followed for the development of CRC. Colorectal cancers diagnosed outside of the screening program (including CRCs in participants who did not return to screening, FIT interval cancers, post-colonoscopy CRCs, and CRCs detected at surveillance colonoscopy) were identified through linkage with the Dutch Comprehensive Cancer Centre (www.iknl.nl), which was up to date until March 2015.

Statistical analysis

Descriptive data were reported as proportions or means with standard deviation (±SD). For non-normally distributed data, the median and interquartile range (IQR) were given. To investigate the role of a negative baseline FIT value, $_{f}$Hb concentrations were divided into six categories; 0, >0-2, ≥2-4, ≥4-6, ≥6-8 and ≥8-10. Per category the cumulative incidence of AN over four rounds was calculated using life tables and curves. Patients were censored at the end of follow up if the event (i.e. AN) had not occurred.

A Cox proportional hazard regression analysis was performed to calculate hazard ratios (HRs), including 95% confidence intervals, to identify factors associated with the development of AN. These factors included age, gender, socioeconomic status (SES), and baseline $_{f}$Hb. The date of the baseline FIT_{bco} was defined as time 0. Only $_{f}$Hb concentrations below 10 µg Hb/g feces were used in these analyses. Factors with a p-value of < 0.10 in univariate analysis, were included in a multivariate model. Interaction terms were also evaluated in the multivariate model. A two-sided P value of < 0.05 was considered to be statistically significant. Analyses were performed using SPSS 21.0 statistics software (IBM Corp., Armonk, New York, USA).

The prediction of an event (i.e. AN or CRC) given the outcome of $_{f}$Hb concentration of a FIT_{bco} at any round (i.e. visit) was analyzed applying a logistic regression technique allowing for multiple repeated measurements per subject using SAS PROC GENMOD with the REPEATED statement with an independent variance assumption.[16] Each visit was hereby associated with the event. The analysis

included adjustment for age at first round (or age at time of FIT screening), sex, and time of round in relation to event or last follow-up. Interactions between gender and age as well as non-linearity of age and $_f$Hb concentration were tested using both pre-specified groups and polynomial regression. Factors with a P value < 0.10 in univariate analysis, were included in the multivariate model. A two-sided P value < 0.05 was considered to be statistically significant. In addition, all analyses were adjusted for multiple testing. Using the results of logistic regression analyses heat plots were generated to depict the risk of AN after two FIT_{bco}. Analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA). All authors had access to the study data and reviewed and approved the final manuscript.

RESULTS

Baseline characteristics

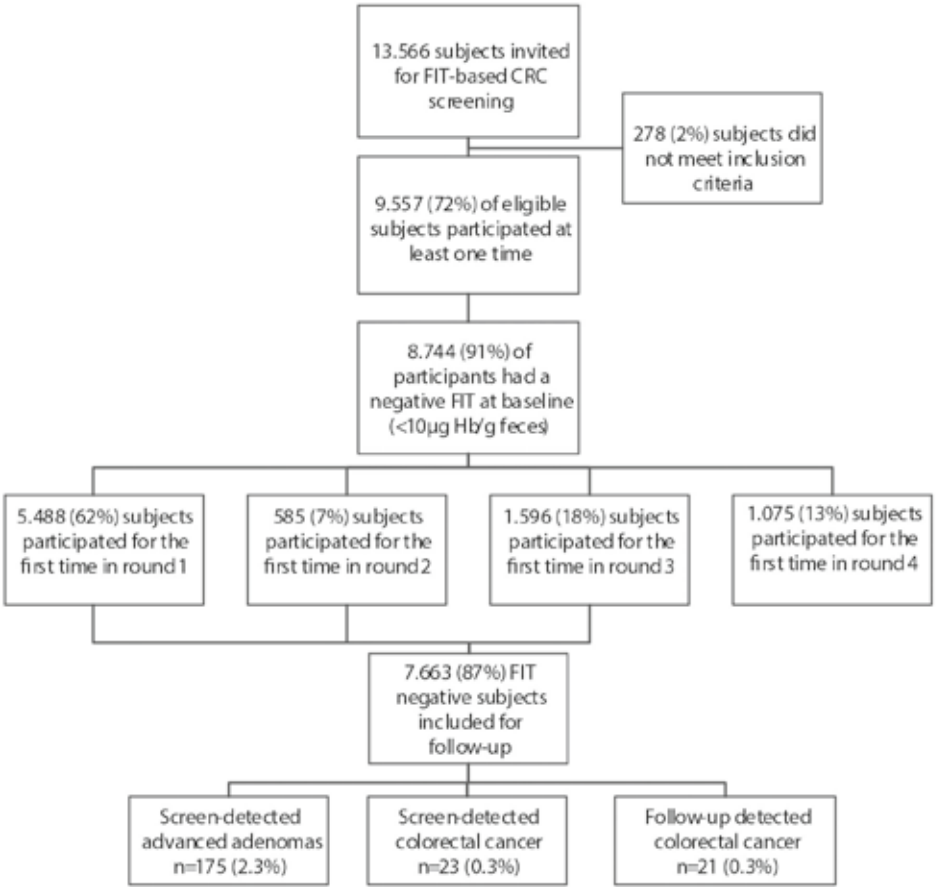
Over four rounds of biennial FIT screening, a total 13,566 subjects were invited of whom 278 (2.0%) did not meet the inclusion criteria. Out of the 13,288 remaining subjects 9,557 (71.9%) participated at least once (Figure 5.1). Out of these participants, 8,744 (91.4%) had a FIT_{bco} at first participation and were included for analysis. Median age was 58 years (IQR 52-64 years) at baseline, with 47·1% of the screenees being male. Overall, 3,172 (36.3%) FIT_{bco} subjects participated in all four rounds, 1,235 (14.3%) in three rounds, 2,254 (25.8%) in two rounds, and 2·059 (23.6%) in one round. Median time between screening rounds was 23 months (IQR 22 - 24 months).

Role of baseline $_f$Hb in predicting risk of advanced neoplasia and colorectal caner

The majority of subjects (62.7%) participated for the first time in round one, the remaining subjects participated in round two (6.6%), round three (18.3%), and round four (12.3%). Because follow-up ended after the fourth round, only baseline $_f$Hb of subjects participating in one of the first three rounds (n=7,663) were included for survival analyses. Median follow up was 4.7 years (IQR 2.0-6.1 years). Over the following rounds, 821 (10.7%) participants with a baseline FIT_{bco} had a positive FIT, of whom 91.8% underwent colonoscopy (n=754). During follow-up 221 (3.0%) cases were diagnosed with AN; 175 with advanced adenomas, 24 with screen-detected CRC, and 22 with CRC detected during follow-up, of which 8 were FIT interval cancers (Figure 5.1).

Cumulative incidences of AN per category of $_f$Hb are shown in Figure 5.2. After 5 years of follow-up, screenees with a baseline $_f$Hb concentration between 8 and

Figure 5.1 Flowchart of study design and outcomes.



10 µg Hb/g had an 8-fold higher cumulative incidence of AN than screenees with a baseline $_{f}$Hb concentration of 0 µg Hb/g ($p < 0{\cdot}001$). After 8 years of follow-up the cumulative incidence increased with 28% compared to screenees with a baseline $_{f}$Hb concentration of 0 µg Hb/g feces ($p<0.001$).

Cox proportional hazard regression analysis showed baseline $_{f}$Hb concentration was associated with the hazard of developing AN in multivariate analysis (Table 5.1). Compared to screenees with a baseline $_{f}$Hb concentration of 0 µg Hb/g feces, HRs increased from 1.2 (95% CI 0.9-1.7) for $_{f}$Hb concentrations between >0 and 2 µg Hb/g, to 8.2 (95% CI 4.5-15.0) for $_{f}$Hb concentrations of ≥8 to 10 µg Hb/g ($p<0.001$).

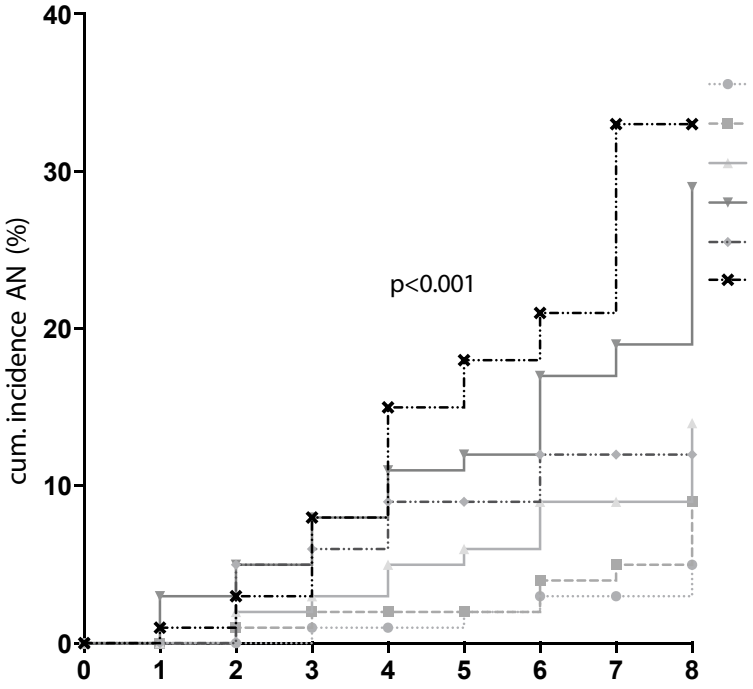
Table 5.1 Cox-regression analysis of baseline $_{f}$Hb concentration for hazard of the detection of advanced neoplasia during follow-up.

	Advanced neoplasia			
	Univariate		Multivariate	
	HR	**95% CI**	**HR**	**95%CI**
Gender (male)	1·7*	1.3-2·3	1·6*	1·2-2·1
Age (years)	1·1*	1·0-1·1	1·1*	1·0-1·1
Baseline $_{f}$Hb concentration				
0 µg Hb/g	Ref.*		Ref.*	
>0-2 µg Hb/g	1·3	0·9-1·8	1·2	0·9-1·7
≥2-4 µg Hb/g	2·9	1·8-4·6	2·8	1·7-4·4
≥4-6 µg Hb/g	6·5	4·2-10·2	5·7	3·7-8·9
≥6-8 µg Hb/g	4·5	2·1-9·3	4·2	2·1-8·7
≥8-10 µg Hb/g	8·9	4·9-16·4	8·2	4·5-15·0
Socioeconomic status				
High	Ref.			
Average	1·0	0·7-1·3		
Low	0·6	0·4-1·0		

* $p<0·001$

$_{f}$Hb; fecal hemoglobin concentration, HR; hazard ratio, CI; confidence interval, conc; concentration, ref; reference category.

Figure 5.2 Life table and curve for advanced neoplasia by $_{f}$Hb level per 2µg Hb/g.
This figure shows that the effect on cumulative incidence of AN of baseline FIT is most prominent for $_{f}$Hb between 4 and 10 µg Hb/g.



	1	2	3	4	5	6	7	8	Time (years)
$_{f}$Hb (µg Hb/g)									
0	4,927	4,185	3,639	2,852	2,410	2,090	1,485	726	Subjects at risk
>0-2	1,874	1,672	1,587	1,427	1,291	1,177	470	94	Subjects at risk
≥2-4	436	376	333	286	247	220	101	34	Subjects at risk
≥4-6	214	171	151	125	105	94	54	16	Subjects at risk
≥6-8	106	87	78	65	56	47	18	4	Subjects at risk
≥8-10	78	70	58	44	35	29	18	7	Subjects at risk

Role of $_{f}$Hb concentrations of consecutive FITs in predicting risk of advanced neoplasia

To further explore the use of $_{f}$Hb of consecutive FITs, $_{f}$Hb concentrations of all screenees that had at least two consecutive $FITs_{bco}$ were analyzed. In multivariate logistic regression, including time between FITs and total follow-up time, several FIT_{bco} combinations were evaluated to determine relative risk of developing AN, these are shown in Table 5.2. ORs remained similar regardless of the sequence of $_{f}$Hb results.

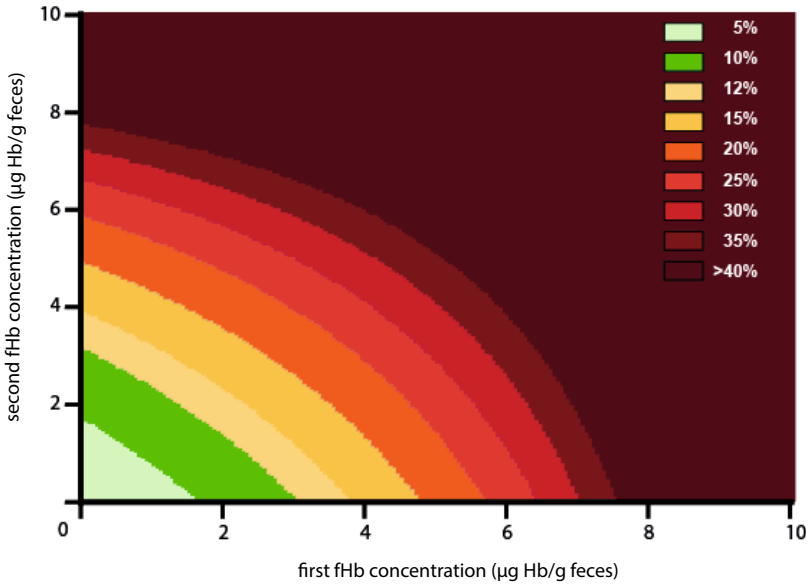
Based on the logistic regression model risks were predicted for all possible combinations of $_{f}$Hb concentrations per 0.5 µg Hb/g feces. Using these risks, heat plots were generated for males and females starting screening from the age of 55 years (Figure 5.3). These heat plots visualize the risk of AN after two consecutive FIT_{bco} at any round according to the combination of $_{f}$Hb concentrations. These heat plots highlight the increased risk of screenees with $_{f}$Hb concentrations up to 10 µg Hb/g and illustrate two-fold increased risk of AN for men compared to women.

Table 5.2 Multivariate logistic regression analysis evaluating various hypothetical combinations of $_{f}$Hb concentration of two previously negative FITs and odds ratio (OR) of developing advanced neoplasia during follow-up.

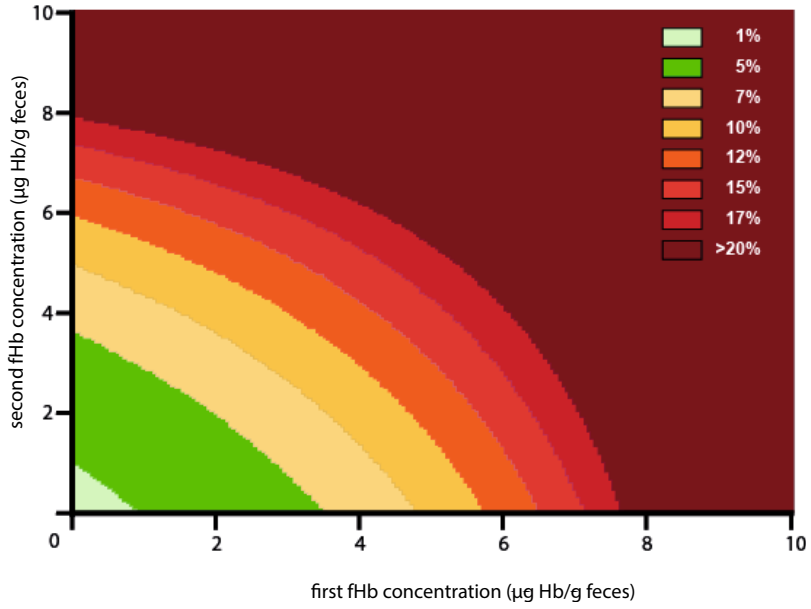
	Advanced neoplasia		
	Multivariate		
	OR	**95% CI**	**p-value**
Gender (female)	2·1	1·3-3·2	0·001
Age (years)	1·0	1·0-1·1	0·04
Combination of first and second $_{f}$Hb concentration			
0 µg Hb/g and 0 µg Hb/g	Ref.		<0·001
1 µg Hb/g and 1 µg Hb/g	1·7	1·5-1·9	
1 µg Hb/g and 5 µg Hb/g	4·4	3·1-6·3	
5 µg Hb/g and 1 µg Hb/g	4·5	3·1-6·6	
5 µg Hb/g and 5 µg Hb/g	7·8	4·6-13·3	
1 µg Hb/g and 8 µg Hb/g	9·0	5·2-15·6	
8 µg Hb/g and 8 µg Hb/g	14·3	4·8-42·3	

Figure 5.3 Heat plot of risk of advanced neoplasia (AN) during further follow up in screenees with two consecutive negative FITs for men (A) and women (B) starting screening at age 55 years. This plot illustrates the increased risk of AN according to $_{f}$Hb concentration. Notable is the two-fold increase for men compared to women.

A



B



DISCUSSION

This study shows that, in a repeated-round FIT-based CRC screening program, fecal hemoglobin concentration below the cut-off is an independent predictor of incident AN. We demonstrated over 8 years of follow-up that screenees with a $_{f}$Hb concentration >0 µg Hb/g to 10 µg Hb/g have a significantly higher cumulative incidence of AN compared to those with a $_{f}$Hb concentration of 0 µg Hb/g. The risk of AN was 8-fold higher for screenees with a $_{f}$Hb concentration of more than 8 µg Hb/g compared to screenees with a $_{f}$Hb concentration of 0 µg Hb/g. We further show that this predictive capacity becomes even stronger with consecutive FIT_{bco} results obtained over repeated rounds.

One other research group has studied the role of FIT concentrations below the cut-off.[12] This Taiwanese study used a cut-off of 20 µg Hb/g feces (OC-Sensor 100 ng Hb/ml buffer), and showed that HRs for AN during follow-up increased with $_{f}$Hb concentration up to 3.4 for subjects with an $_{f}$Hb level of 16-20 µg Hb/g feces. This HR is substantially lower than the HRs found in our study. This can be explained by the fact that the investigators used $_{f}$Hb concentrations between 1-4 µg Hb/g feces as a reference and screenees with a $_{f}$Hb concentration of 0 were not included in multivariate survival analysis. The results of this study were however hampered by the high rate of positive FIT screenees who did not undergo colonoscopy (42%) yet were included in the analyses. While there is no other literature on the use of quantitative FIT_{bco} results, a few studies did look at prior FIT_{bco} results of qualitative tests. An Australian study investigated the use of FIT in a colonoscopy surveillance program and found that subjects with a FIT_{bco} had the lowest risk of AN.[17] A Chinese study compared the number of FIT_{bco} and found no differences in outcome between subjects with one FIT_{bco} versus subjects with three subsequent FIT_{bco}.[18] However, as the $_{f}$Hb concentrations were not reported, results could not be stratified according to $_{f}$Hb concentration, and comparison of these results to our findings is not possible.

To the best of our knowledge, we are the first to explore the use of $_{f}$Hb concentrations of FIT_{bco} over consecutive rounds as a predictive variable for AN in population-based CRC screening. Exploring $_{f}$Hb concentrations over the course of years makes sense, as it has been hypothesized that during the development from adenoma to carcinoma, in the adenoma-carcinoma sequence, adenomas will increasingly bleed. This natural history of adenomas is supported by our findings. Our results are further strengthened by the finding that in screenees who participated in all four rounds $_{f}$Hb increased among those that were diagnosed with AN in the fourth round (Supplementary file 5.1). A similar trend was also

described by the Taiwanese study, demonstrating that median $_{f}$Hb increases over rounds among screenees that are diagnosed with AN in a later round.[12]

Strengths of this study include the analysis over multiple rounds, stratifying for FIT_{bco} levels unto nil µg Hb/g feces. Also, only average risk individuals were included, and the program consisted of true population-based screening, with a consistent screening protocol. This makes these results applicable to all fecal occult blood screening programs worldwide.[19] Our study also has its limitations, such as the limited number of subjects diagnosed with CRC after having participated in three or four rounds. One further issue of this study is that it is susceptible to verification bias, due to the fact that only FIT positive screenees are referred for colonoscopy. As such, the yield of AN could be equally high in screenees with low FIT-values, but these simply do not receive verification by colonoscopy. To partially assess the possibility of verification bias, we performed two additional analysis in which we compared yield of AN only in screenees with a positive FIT during follow-up or screenees with consistent FIT_{bco} (i.e. interval cancer rate; Supplementary File 5.2). Although numbers are small, both analyses consistently showed that the yield of AN was higher in those with higher levels of $_{f}$Hb, similar as our base case analysis. Next, data from our primary colonoscopy screening trail in the same region and time period also showed that FIT_{bco} predicts AN in a single round of FIT (Supplementary file 5.3).[20]These findings suggest that the impact of verification bias is small and corroborate our finding that FIT_{bco} is predictive for future AN. Despite the verification bias, the lack of colonoscopy in FIT negative screenees, is also a strength of our study. If all screenees would undergo colonoscopy, the opportunity to follow FIT_{bco} screenees for the development of AN would be lost.

As more screening programs are being implemented worldwide and FIT is gaining popularity, the use of quantitative FIT should be further explored. Expressing FIT-results not solely as a positive or negative result, but incorporating $_{f}$Hb concentrations in risk prediction models to estimate risk of AN can improve screening efficiency. Our results justify evaluation of screening strategies in which $_{f}$Hb concentrations are used to establish screening intervals. In current practice, a screenee with a FIT_{bco} is re-invited in the next screening round to perform a 'new' FIT. This FIT result is used as a referral criterion for colonoscopy regardless of the $_{f}$Hb concentration measured in the previous round. By neglecting the previous FIT result, an opportunity is lost to use the quantitative information of two FITs for risk stratification. We demonstrate, as depicted by the heat plots, that previous FITs enable identification of those at low risk (e.g. screenees who twice had a result of 0 µg Hb/g) and those at considerable risk. Such heat plots can be of use to visualize

risks of AN for screenees and health care professionals. Identifying those at high risk of AN would possibly decrease interval carcinoma rate. Next, combining $_{f}$Hb with known risk factors could facilitate establishing individual screening intervals. Such individualized screening intervals may increase adherence to screening, as the majority of participants (i.e. those with consecutive $_{f}$Hb concentrations of 0 µg Hb/g feces and a low risk of AN) will then have to be screened less frequently. This in turn could increase screening efficiency. As a consequence, screening and subsequently colonoscopy demand would decrease. The latter is especially important, as many countries are struggling with limited colonoscopy capacity.[8 21]

The results of this prospective FIT-based CRC screening cohort show that $_{f}$Hb concentration of a FIT_{bco} in a first round of screening is an independent predictor for the risk of incident AN. Furthermore, consecutive FIT_{bco} could be used in determining personalized screening strategies. These findings pave the way for the use of $_{f}$Hb concentration in both public health programs as well as clinical decision-making. These results aid in informing patients about the risk of AN after multiple FIT_{bco}, and to alter screening intervals accordingly.

REFERENCES

1. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nat Rev Dis Primers* 2015;**1**(In Press):15065.
2. 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.
3. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. *Nat Rev Clin Oncol* 2013;**10**(3):130-42.
4. van Doorn SC, Stegeman I, Stroobants AK, et al. Fecal immunochemical testing results and characteristics of colonic lesions. *Endoscopy* 2015;**47**(11):1011-7.
5. Digby J, Fraser CG, Carey FA, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. *J Clin Pathol* 2013;**66**(5):415-9.
6. Fraser CG, Mathew CM, McKay K, et al. Automated immunochemical quantitation of haemoglobin in faeces collected on cards for screening for colorectal cancer. *Gut* 2008;**57**(9):1256-60.
7. 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.
8. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut* 2015;**64**(10):1637-49.
9. Kanavos P, Schurer W. The dynamics of colorectal cancer management in 17 countries. *Eur J Health Econ* 2010;**10 Suppl 1**:S115-29.
10. 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**(3):198-205.
11. van Hees F, Zauber AG, van Veldhuizen H, 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.
12. Chen LS, Yen AM, Chiu SY, et al. Baseline faecal occult blood concentration as a predictor of incident colorectal neoplasia: longitudinal follow-up of a Taiwanese population-based colorectal cancer screening cohort. *Lancet Oncol* 2011;**12**(6):551-8.
13. Ciatto S, Martinelli F, Castiglione G, et al. Association of FOBT-assessed faecal Hb content with colonic lesions detected in the Florence screening programme. *Br J Cancer* 2007;**96**(2):218-21.
14. Kapidzic A, Grobbee EJ, Hol L, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. *Am J Gastroenterol* 2014;**109**(8):1257-64.
15. Grobbee EJ, van der Vlugt M, van Vuuren A, et al. Comparison of OC-Sensor and FOB-Gold in Population-Based Colorectal Cancer Screening Based on FIT. *Gastroenterology*;**148**(4):S-160.
16. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;**11**(5):561-70.
17. Lane JM, Chow E, Young GP, et al. Interval fecal immunochemical testing in a colonoscopic surveillance program speeds detection of colorectal neoplasia. *Gastroenterology* 2010;**139**(6):1918-26.
18. Wong MC, Ching JY, Chan VC, et al. Should prior FIT results be incorporated as an additional variable to estimate risk of colorectal neoplasia? A prospective study of 5,813 screening colonoscopies.

PLoS One 2014;**9**(12):e114332.

19. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut* 2015

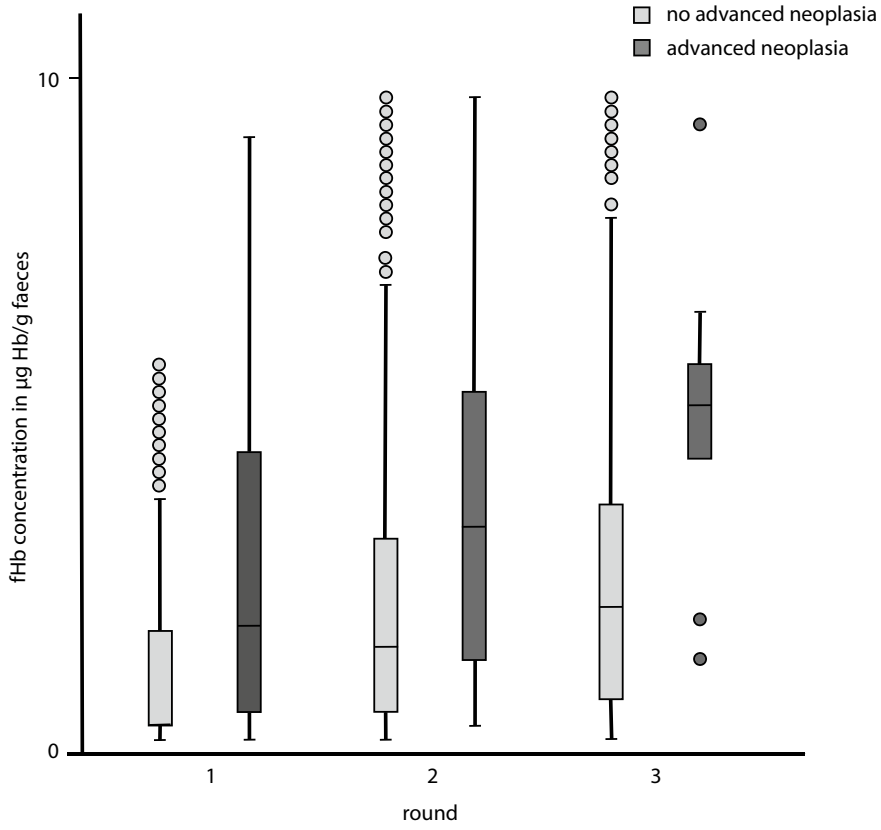
20. de Wijkerslooth TR, de Haan MC, Stoop EM, et al. Study protocol: population screening for colorectal cancer by colonoscopy or CT colonography: a randomized controlled trial. *BMC Gastroenterol* 2010;**10**:47.

21. Digby J, Fraser CG, Carey FA, et al. Interval cancers using a quantitative faecal immunochemical test (FIT) for haemoglobin when colonoscopy capacity is limited. *J Med Screen* 2015

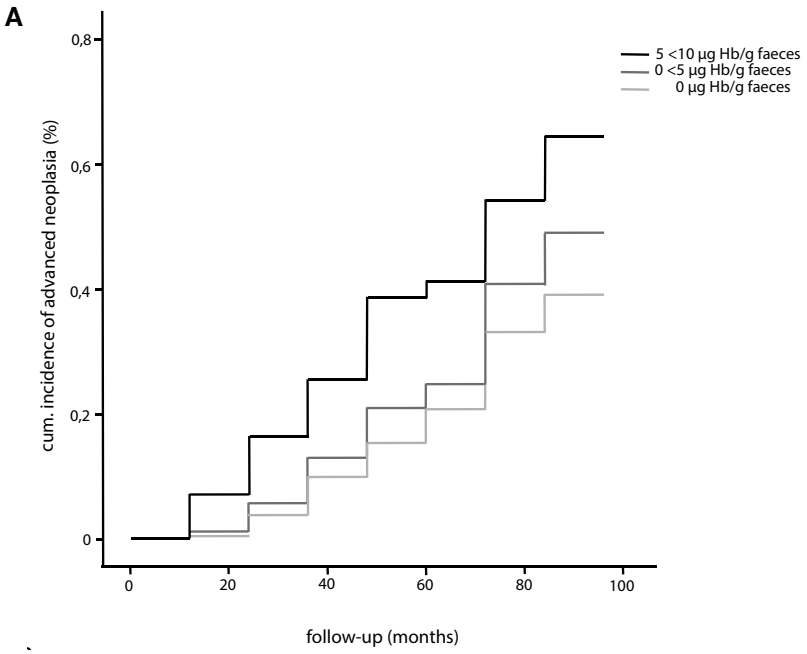
SUPPLEMENTARY FILES

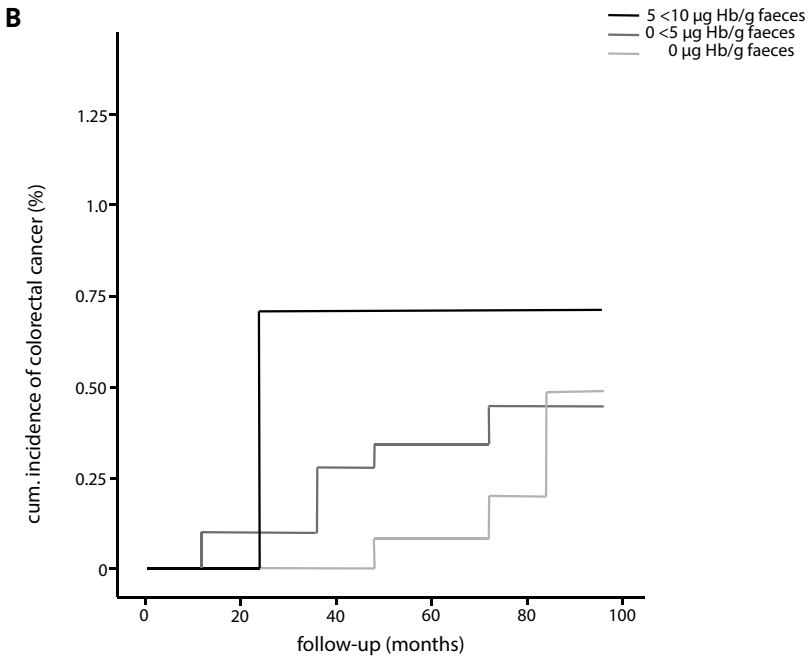
Supplementary file 5.1 Median fecal Hb of subjects with FIT_{bco} of > 0µg Hb/g that participated in all 4 rounds.

Results are presented by occurrence of advanced neoplasia (AN) detected in the 4th round. This figure shows an increase in median $_{f}$Hb concentration over consecutive rounds in subjects who had AN detected in the fourth round, indicating that those screenees that bleed for other reasons than AN, show less of an increase in $_{f}$Hb over rounds.

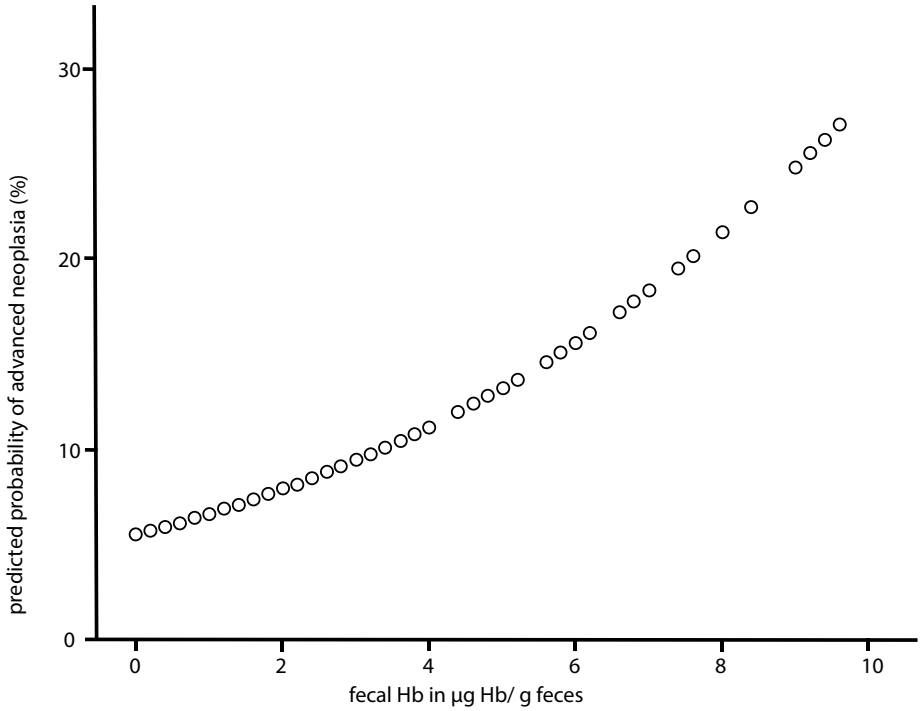


Supplementary file 5.2 Life table of FIT_{bco} per 5 µg Hb/g in those who had a positive FIT during follow-up (p<0.001; **A**) and those who never had a positive FIT in one of the four rounds (p=0.034;**B**).





Supplementary file 5.3 Predicted probability of advanced neoplasia (AN) of all subjects with fHb <10 µg Hb/g from previously published data.[20]



PART THREE

Quality in screening and colonoscopy

Chapter 6	Variable quality and readability of patient-oriented websites on colorectal cancer screening *Clinical Gastroenterology and Hepatology. 2016 Jul 9. (Epub ahead of print)*
Chapter 7	Plenary feedback on inter-hospital differences to improve variation in quality of colonoscopy *Manuscript in preparation*
Chapter 8	The appropriateness of surveillance colonoscopy intervals after polypectomy *Canadian Journal of Gastroenterology. 2013 Jan;27(1):33-8.*

CHAPTER 6

Variable quality and readability of patient-oriented websites on colorectal cancer screening

E.H. Schreuders
E.J. Grobbee
M.C.W. Spaander
E.J. Kuipers
S.J.O. Veldhuyzen van Zanten

Clinical Gastroenterology and Hepatology. 2016 Jul 9. (Epub ahead of print)

ABSTRACT

Background & Aims

The efficacy of colorectal cancer (CRC) screening is dependent on participation and subsequent adherence to surveillance. The internet is increasingly used for health information and important to support decision-making. We evaluated the accuracy, quality, and readability of online information on CRC screening and surveillance.

Methods

A Website Accuracy Score and Polyp Score were developed, which award points for various aspects of CRC screening and surveillance. Websites were also evaluated using validated internet quality instruments (Global Quality Score, LIDA and DISCERN), and reading scores. Two raters independently assessed the top-30 websites appearing in Google.com™. Portals, duplicates and news articles were excluded.

Results

Twenty websites were included. The mean Website Accuracy Score was 26 out of 44 (range 9-41). Websites with the highest scores were www.cancer.org, www.bowelcanceraustralia.org and www.uptodate.com. Median Polyp Score was three out of ten. The median Global Quality Score was three out of five (range 2-5). The median LIDA overall score was 74% and median DISCERN score was 45, both indicating moderate quality. The mean Flesch-Kincaid Grade Level was 11th grade rating the websites as difficult to read, 30% had a reading level acceptable for the general public (Flesch Reading Ease>60). There was no correlation between the Google rank and Website Accuracy Score (r_s=-0.31; p=0.18).

Conclusions

There is marked variation in quality and readability of websites on CRC screening. Most websites do not address polyp surveillance. The poor correlation between quality and Google ranking suggests that screenees will miss out on high-quality websites using standard search strategies.

INTRODUCTION

Screening is effective in reducing the burden of colorectal cancer (CRC) and many countries have implemented CRC screening programs.[1 2] The success of CRC prevention is highly dependent on participation in the screening program. Initial participation and subsequent adherence to surveillance can be influenced by enhanced knowledge about CRC screening and colonoscopy outcome.[3 4] As more screening programs are implemented worldwide, providing adequate patient oriented information is increasingly important. Most organized screening programs approach individuals for screening on a voluntary basis without personal contact with a health professional.[2 5] Accordingly these individuals may search for additional information on screening themselves.

The internet is widely regarded as an important channel of health information.[6 7] In Western countries, more than half of the population uses a smartphone allowing instant and rapid access to the worldwide web.[8] However, few regulations control the information that individuals or organizations list on their web sites. A systematic review reported that 70% of studies identified quality issues with health- and disease-focused internet websites.[9] Since the efficacy of a CRC screening program is dependent on informed participation, assessing the availability and quality of online information aimed at screenees is of crucial importance. Therefore, the aim of this study was to rate quality, accuracy and readability of web-based information on CRC screening from a screenee perspective.

METHODS

Internet Search Strategy

Websites were identified by searching the World Wide Web with Google.com™, the most frequently used Internet search engine.[10] The search was performed with English settings, with location tracking and search activity history switched off so search results were not influenced by location or past searches. Searches were carried out in 2014, 2015 and 2016. The following search terms were used: "colorectal cancer screening" OR "bowel cancer screening" OR "colon cancer screening" (quotations included). The search terms used reflect the most searched terms listed in the statistics provided via Google Trends (Supplementary file 3.1).

It is known that internet searchers do not typically view more than a few search hits and usually choose one of the first results displayed by the search engine.[11] We

therefore decided to examine the first 30 hits, corresponding with the first three pages of Google searches.

Inclusion and exclusion criteria

English websites were included only if the main part of the site dealt with educational information about CRC-screening. Websites that merely contained portal links to other sites were excluded, as were duplicate websites, news articles and sites containing irrelevant information (e.g. advertising, retail sites, or patient fora).

Accuracy assessment

The variability and accuracy of the information provided by each website on key facts about CRC screening and surveillance was investigated. For this purpose a Website Accuracy Score specific for CRC screening was developed (Table 6.1). In addition a separate Polyp Score for colorectal polyps was developed to assess information on important aspects of polyps, colonoscopy outcome and surveillance guidelines (Table 6.2). The Website Accuracy Score and Polyp Score consist of a list of key items deemed relevant for CRC screening and surveillance. They were generated through evaluation of the literature and discussions with key stakeholders. The Website Accuracy Score and Polyp Score went through five iterations and were pretested twice prior to its final use using a random selection of websites. The range of scores was 0 to 44 for the Website Accuracy Score and 0 to 10 for the Polyp Score. If a website did not discuss or name an item of the Website Accuracy Score or Polyp Score, zero points were awarded for that item. Items had to be clearly presented on the website; the search function of the website was not used to locate this information.

Quality assessment

In addition to Website Accuracy Score and the Polyp Score, a selection of validated scores was used to assess the website quality and reliability. The overall quality of each website was rated using the *Global Quality Score*. This is a previously validated five-point Likert scale to rate the overall quality of a website (Table 6.3).[12 13] It incorporates the accessibility of the information within the website, the quality of this information, the overall flow of information, and how useful the website reviewer thinks the particular website would be to a screenee. The Global Quality Score was assigned by the reviewer after evaluating the entire website.

Table 6.3. Global Quality Score criteria used to score websites on CRC screening

Score	Global Quality Score Description
1	Poor quality, poor flow of the site, most information missing, not at all useful for patients.
2	Generally poor quality and poor flow, some information listed but many important topics missing, of very limited use to patients.
3	Moderate quality, suboptimal flow, some important information is adequately discussed but other information poorly discussed, somewhat useful for patients.
4	Good quality and generally good flow, most of the relevant information is listed, but some topics not covered, useful for patients.
5	Excellent quality and excellent flow, very useful for patients.

CRC; Colorectal cancer

The *LIDA instrument* is a validated question based instrument, assessing the overall score (0-96), accessibility (0-54), usability (0-12) and reliability (0-30) of healthcare websites. The scores are reported as percentages of the maximum score, overall scores >90% represent good results and <50% represent poor results. The online LIDA instrument was used for this study.[14]

The *DISCERN tool* is a validated 16-item questionnaire to rate the quality of written information on treatment choices for a health problem.[15 16] The first 8 questions address reliability, dependability and trustworthiness of a website, the next 7 questions focus on quality of information on treatment choices and the last question addresses the overall quality of the site. Each question is rated on a 5-point scale with a maximum score of 80. Questions were answered as if participation to CRC screening was the treatment choice. The total quality of each website was classified as high (≥65 points), moderate (33–64 points), or low (16–32 points).

The amount of advertisements on each website was scored as none, little, average, or many and agreed through discussion by the two reviewers.

Readability assessment

Readability, referring to the reading difficulty based on word and sentence length, was assessed by the use of two readability scores. The *Flesch Reading Ease Score (FRE)* assigns a value between 0 and 100 whereby a higher value represents a greater ease of reading. A section with a score of 90–100 is considered to be very easily understood, >60 is an acceptable level of difficulty for the general

public and below 30 is considered very difficult to read.[17] The *Flesch-Kincaid Grade Level (FKG)* uses the same input variables as the FRE score and outputs a US school grade indicating the average school grade able to read the text .[17] The American Medical Association Foundation states that health-related materials for patients should be written at a level appropriate for those in the 6th grade or below.[18] The FRE score and FKG score were calculated using the Microsoft Word 2007 program. A random 100 word sample of text was extracted from each website and pasted into the program by both reviewers independently.

Statistical analyses

The website assessment was performed by two independent raters (EHS, EJG). For Website Accuracy Score assessment, any difference in score between reviewers was resolved through discussion and by re-review of the website by both reviewers together to generate a single score for each website. Consensus in case of disagreement was achieved through discussion with a third reviewer (SvZ). For other quality parameters, the mean score of both website raters was used. Correlations between different quality parameters were analyzed using the Spearman rank-correlation coefficient because of non-normality of the data. Statistical tests were performed with the use of IBM SPSS software, version 21.0 and Graphpad Prism 5. A two-sided p-value <0.05 was considered significant. All authors had access to the study data and reviewed and approved the final manuscript.

RESULTS

The search

The Google search was carried out on April 9th, 2014 and resulted in over 2,000,000 hits. The first 30 results were evaluated of which 20 websites were included. Two portal websites leading to another site, one duplicate site, one website with information on insurance reimbursement, three news articles, and three guidelines and medical articles clearly aimed at health professionals were excluded. All websites were accessed between April 2014 and June 2014. Additional Google searches were carried out on August 7th, 2015 and February 22nd, 2016 to evaluate possible changes in Google rank position. Most websites were published by a professional medical society (35%) or a governmental organization (30%) (Figure 6.2). Almost half of the websites were from the United States (45%), others from the United Kingdom (25%), Canada (20%) and Australia (10%).

Table 6.1. Colorectal cancer screening specific Website Accuracy Score components and percentage of websites that were awarded points for these items.

Website information components (maximum 44 points)	**Websites N (%)**
CRC general information	
Description of the colon/bowel/large intestine	15 (75)
Image of the anatomy of the intestines	15 (75)
Explanation of polyp as precursor of colorectal cancer	17 (85)
Development of a polyp into malignancy is a slow process (takes years)	8 (40)
Colorectal cancer can be prevented by removing precancerous polyps/adenomas	15 (75)
Causes of CRC	
Risk factors	
Unknown	3 (15)
Age (> 50)	13 (65)
Gender (male)	0 (0)
History of previous polyps	15 (75)
Family history of colorectal cancer	17 (85)
Hereditary / Familial adenomatous polyposis / Lynch syndrome	14 (70)
Life style (2)	
Unhealthy lifestyle (general)	
Unhealthy diet (low fiber, high fat, red meat)	
Smoking	
Alcohol	1 (5)
Obesity	13 (65)
Mentions 1-2 life style factors: 1 point	
Mentions 3 or more life style factors: 2 points	
Symptoms of colonic polyps / CRC	
Most polyps are asymptomatic	11 (55)
Mentions symptom(s) such as:	13 (65)
blood in stool/ rectal bleeding	
change in bowel habit	
unexplained weight loss	
tenesmus (false urge)	
Recommendation to contact medical doctor in case of symptoms	11 (55)

Website information components (maximum 44 points)	Websites N (%)
Screening for CRC	
Mentions that there are different methods of screening	16 (80)
The detection and removal of polyps is main purpose of the screening program for colorectal cancer	15 (75)
Mentions that not all tests have same accuracy	7 (35)
Mentions that not all tests have same patient burden	7 (35)
Colonoscopy is gold standard / most accurate for diagnosing polyps	7 (35)
Colonoscopy	20 (100)
+ explanation of procedure	17 (85)
+ explanation risks (bleeding and perforation are mentioned)	14 (70)
+ explanation polypectomy	13 (65)
+ explanation bowel preparation	13 (65)
Mentions flexible sigmoidoscopy	15 (75)
+ explanation procedure	13 (65)
+ explanation risks	8 (40)
Mentions FOBT (immunochemical or guaiac)	20 (100)
+ explanation procedure	16 (80)
+ has to be repeated every 1-2 years	13 (65)
+ stresses importance of repeated screening	7 (35)
+ explains possibility of false positive/negative results	11 (55)
Mentions barium enema	9 (45)
+ poor detection of (pre)cancer	3 (15)
Mentions CT colonography	11 (55)
+ explanation procedure	10 (50)
+ explanation risks	8 (40)
Mentions that all tests, when positive, need to be followed by colonoscopy	9 (45)
Mentions surveillance after colonoscopy in case of adenomas	3 (15)
Mentions that frequency of screening is different per test	7 (35)
Describes possibility of interval carcinomas (CRC after negative test)	4 (20)
Describes limitations of screening such as overdiagnosis and overtreatment	5 (25)

Maximum number of points is 1 per item, unless otherwise specified.
CRC; Colorectal cancer, CT; Computed Tomography, FOBT; Fecal Occult Blood Test.

Table 6.2 Polyp Score items and percentage of websites that were awarded points for these items

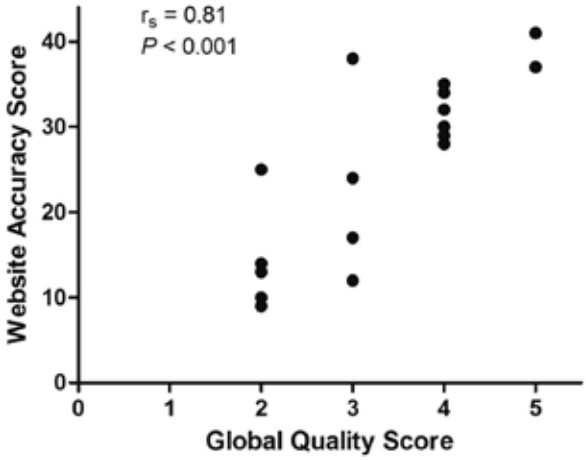
Website information components (maximum 10 points)	Websites, n (%)
Description of what a polyp is; growth/mushroom/lump in the lining of the large bowel	15 (75)
Image of a polyp	9 (45)
Prevalence of people with polyps in population	6 (30)
Explains that there are different types of polyps	9 (45)
Explains that not all polyps have an equal risk of turning in to colon cancer	10 (50)
Explains differences between adenoma and hyperplastic polyp	4 (20)
Mentions that some polyp characteristics have a higher risk of malignant degeneration; i.e. histological findings (villous aspect)	3 (15)
Polyp size as risk factor	3 (15)
Influence of degree of cleanliness of bowel on polyp detection	2 (10)
Explains surveillances intervals after polypectomy	2 (10)

Maximum number of points is 1 per item.

Accuracy and quality of website information

The mean Website Accuracy Score was 26 (range 9-41). Most websites contained general information on CRC screening, but description and risk of different screening modalities and limitations of screening were not always captured (Table 6.1). The median Global Quality Score was 3 (range 2-5). This score indicates that the quality of information of most websites was moderate. In many sites, some information was adequately discussed, while other parts of information were missing and the overall flow of information was suboptimal. There was a strong positive correlation between the Website Accuracy Score and the Global Quality Score with a Spearman's rho (r_s) of 0.81 ($p<0.001$) (Figure 6.1). The median Polyp Score was three (range 0-10, Table 6.2). The Polyp Score correlated positively with the Global Quality Score ($r_s=0.81$; $p<0.001$). The median LIDA overall score was 74% (Interquartile range, IQR 11). The median LIDA score for accessibility was 88% (IQR 8), for usability 63% (IQR 22), and for reliability 52% (IQR 26). The median DISCERN score was 45 (IQR 20) indicating moderate quality. Ten percent of websites (2/20) were classified by DISCERN as high quality, 80% (16/20) as moderate and 10% (2/20) as low quality. Both the validated LIDA and DISCERN had a moderate correlation with the Website Accuracy Score; $r_s=0.45$ ($p<0.05$) and $r_s=0.66$ ($p<0.01$) respectively. There was no correlation between the Google ranks and the Website Accuracy Score ($r_s=-0.31$; $p=0.18$, $r_s=-0.47$; $p=0.08$ and $r_s=-0.31$; $p=0.25$ for the 2014, 2015 and 2016 search respectively).

Figure 6.1 Relationship of the Global Quality Score and the Website Accuracy Score used to evaluate colorectal cancer screening websites



r_s; Spearman's rho.

Table 6.4 lists the top five websites as rated by CRC screening specific Website Accuracy Score and other evaluations of website quality. The complete scores per website are published in a supplementary table (Supplementary file 3.2). Eight websites had initial inter-rater Website Accuracy Score differences of ≥8. Differences in scoring of Website Accuracy Score or Polyp Score between reviewers were due to oversight or differences in interpretation.

Table 6.4 Top 5 websites as ranked by the Website Accuracy Score with the corresponding Polyp Score, quality scores, reading scores and Google rank positions

Website	Accuracy		Quality			Readability		Google rank		
	WAS	PS	GQS	DISCERN	LIDA	FRE	FKG	2014	2015	2016
www.cancer.org	41	5	5	65	67%	62	9th	6	3	2
www.bowelcancer australia.org	38	2	3	35	58%	58	10th	5	X	29
www.uptodate.com	37	6	5	69	85%	28	14th	27	13	16
www.macmillan. org.uk	35	5	4	49	69%	48	11th	19	X	X
www.nlm.nih.gov/ medlineplus	34	6	4	57	81%	59	8th	4	4	4

FRE; Flesch Reading Ease score, FKG; Flesch-Kincaid Grade Level, GQS; Global Quality Score, PS; Polyp Score, WAS; Website Accuracy Score, X; Not in the first 30 Google results.

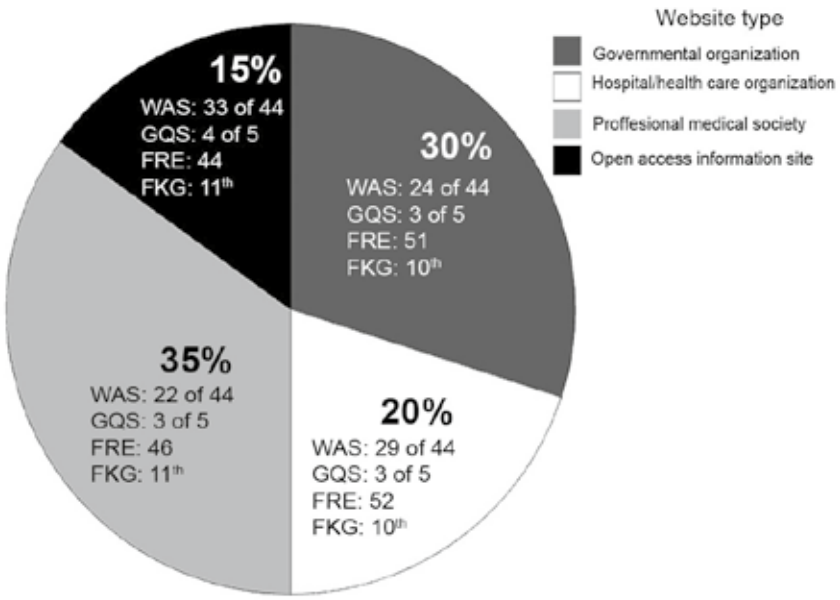
Readability of websites

The mean Flesch Reading Ease Score was 48 (range 27–76), 30% of the websites had a reading level acceptable for the general public defined by a Flesch Reading Ease Score of >60. The mean Flesch-Kincaid Gradel Level was 11 (SD ±2.2, range 5–16), indicating that the text would be understandable to an average 11th grade US student. The reading level of healthcare and governmental websites was the easiest, whereas the reading level of open access information sites was the most difficult (Figure 6.2).

Advertisements

When assessing the amount of advertisement, 16 (80%) websites contained none, two (10%) contained a moderate amount of advertisement, and two (10%) websites contained many advertisements. The latter two were open access websites. Websites published by governmental organizations contained no advertisements.

Figure 6.2 Mean Website Accuracy Score, Global Quality Score and reading scores per website type.



FRE; Flesch Reading Ease score, FKG; Flesch-Kincaid Grade Level, GQS; Global Quality Score, WAS; Website Accuracy Score.

DISCUSSION

This study shows that there is marked variation in accuracy, quality and readability of information on CRC screening websites and that most websites do not address polyp surveillance. The best five websites as ranked by the Website Accuracy Score are www.cancer.org; www.bowelcanceraustralia.org; www.uptodate.com; www.macmillan.org.uk; and www.nlm.nih.gov/medlineplus. Their corresponding Google rank positions varied over time and some of these websites will be missed by standard Google searches (Table 6.4).

The poor correlation between website accuracy and Google ranking is especially concerning given the fact that Google is a prominent search engine.[10] Internet users often do not go beyond the first page of a search, which can result in missing websites that provide high quality information. This problem has been identified before.[12 13 19]

Even though surveillance after colonoscopy, especially if adenomatous polyps were found, is important for CRC screening to reach its maximal efficacy, it was only mentioned in 15% of the websites. Surveillance intervals are based on findings during colonoscopy.[20] However, clear and easy to understand information on how findings during a screening colonoscopy, i.e. adenomatous polyps, determine the follow-up surveillance recommendations was lacking in most sites. This is reflected in the low overall median Polyp Score (3 out of 10) and the fact that only two websites (10%) described the actual surveillance intervals. This is an important information gap since adherence to surveillance is influenced by enhanced knowledge.[4] Previous studies have shown that patients may not be sufficiently aware of important endoscopic findings and the consequences this has for subsequent surveillance recommendations. [3 21] Understanding the need of surveillance likely will motivate participants to adhere to surveillance recommendations.

The reading difficulty of most websites was far above the required standard. Only 5% of the websites met the recommended level by the American Medical Association Foundation of 6th grade or below.[18] This suggests that most websites are too difficult for the average reader and this may result in misunderstanding of information. Other studies evaluating patient information websites also documented that the required reading levels were high and above the recommended 6th grade level.[22-24] Our study showed that commercially funded websites were more difficult to read than governmental websites. This is in accordance with previous literature.[22]

When evaluating the Website Accuracy Scores, it became apparent that most websites only focused on the predominant screening test used in the country where the website originated, and did not provide information on other options for CRC screening. It is debatable whether it is necessary to inform screenees about all possible screening tests that are available.[25] However, providing information that several different options exist may help individuals, who are interested in screening, to make an informed decision.[25 26] Colonoscopy and guaiac or immunochemical Fecal Occult Blood Test (FOBT) were described in all websites in detail. However, not all websites stressed the importance of the need for repeated screening when FOBT is used. This in spite of strong evidence that repeating stool testing at regular intervals is of paramount importance for FOBT-based screening to be effective in the long term.[1] Only 20% of the websites mentioned the possibility of the occurrence of interval carcinomas. This may in part be explained by the fact that this aspect of CRC screening has only gained a lot of attention during the last few years. However, not mentioning potential limitations of screening may stand in the way of informed decision-making.[27]

A strength of this study is that the Website Accuracy Score and Polyp Score are CRC screening specific evaluation tools. These content specific outcome measures showed moderate to strong correlation with the validated generic outcome measures of Global Quality Score, LIDA and DISCERN. This provides further evidence that the use of these CRC screening specific outcome measures provide meaningful and relevant information. The advantage of the Global Quality Score over LIDA and DISCERN is that it is short and easy to perform. We believe that the Global Quality Score is a good score for overall flow and ease of use of any website providing health information.

This is to our knowledge the first review that systematically assessed the quality, accuracy and readability of patient-oriented websites on CRC screening as well as polyp surveillance. Previous studies have reported on the quality of web-based information regarding CRC surgery or treatment but none were systematic reviews of existing websites.[23 28 29] Two other publications evaluated CRC screening websites but these did not include detailed information on polyps and surveillance.[30 31] An American study focused on the readability and suitability of 12 CRC screening websites.[31] However, these sites were self-chosen by the author. Another brief review examined five chosen websites and evaluated their content and usability.[30] In both of these publications no apparent selection criteria for quality were used. Most of the listed websites did not appear in our original 2014 search results, nor in the first three Google pages assessed in 2015 and 2016.

Our study has some limitations. Both the Website Accuracy Score and Polyp

Score were not validated separately prior to use on the selected websites. However, the good correlation with other previously validated quality instruments suggests adequate content validity. We only searched using English search terms thus only English websites were retrieved. Another possible limitation is the fact that quotation based search terms were used which require words to appear together in retrieved websites. Omission of quotation marks when searches are done could lead to different results.

The internet is increasingly used by consumers to find relevant health information. There is evidence that experience and knowledge of Internet use has a significant impact on the uptake of CRC screening.[32] Furthermore, the credibility of cancer-related information on the internet is associated with population compliance with CRC screening, indicating the relevance of this study.[4 32] We believe health care providers interested in developing websites on CRC screening, for example for their own institutions, can use our approach to evaluate the quality and readability of provided information to develop the content of the site they are creating. Alternatively they can provide health care consumers to several of the high quality websites, listed in Table 6.4, that we identified.

Physicians should be aware of the limitation of Google searching for CRC screening. Our study may be helpful in that regard as it provides a list of those websites which provide the highest quality information on CRC screening. However, it is important to remember that the internet is "alive" and that quality of websites may change over time or that new high quality websites may be developed.

In conclusion, our study showed that there is marked variation in overall quality of web-based patient information on CRC screening. Most websites lack important information regarding polyps and their importance for future follow up surveillance colonoscopies. Several high quality websites do exist but poor correlation with Google ranking suggests that these websites may be missed. High quality and readable websites are essential to provide patients with reliable information to make informed decisions on CRC screening and surveillance participation and to optimize efficacy.

REFERENCES

1. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol 2013;**10**(3):130-42.
2. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.
3. Kumaravel V, Heald B, Lopez R, et al. Patients Do Not Recall Important Details About Polyps, Required for Colorectal Cancer Prevention. Clinical Gastroenterology and Hepatology 2013;**11**(5):543-47.e2.
4. Wee CC, McCarthy EP, Phillips RS. Factors associated with colon cancer screening: the role of patient factors and physician counseling. Prev Med 2005;**41**(1):23-9.
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:gutjnl-2014-308074.
6. Webb TL, Joseph J, Yardley L, et al. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;**12**(1):e4.
7. Zulman DM, Kirch M, Zheng K, et al. Trust in the internet as a health resource among older adults: analysis of data from a nationally representative survey. J Med Internet Res 2011;**13**(1):e19.
8. eMarketer. Worldwide Smartphone Usage to Grow 25% in 2014. Secondary Worldwide Smartphone Usage to Grow 25% in 2014 Jun 11, 2014 2014. http://www.emarketer.com/Article/Worldwide-Smartphone-Usage-Grow-25-2014/1010920.
9. Eysenbach G, Powell J, Kuss O, et al. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;**287**(20):2691-700.
10. comScore. qSearch 2015 U.S. Desktop Search Engine Rankings. Secondary qSearch 2015 U.S. Desktop Search Engine Rankings January 20, 2016 2016. https://www.comscore.com/dut/Insights/Rankings/comScore-Releases-December-2015-US-Desktop-Search-Engine-Rankings.
11. Eysenbach G, Kohler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002;**324**(7337):573-7.
12. Bernard A, Langille M, Hughes S, et al. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007;**102**(9):2070-7.
13. Langille M, Bernard A, Rodgers C, et al. Systematic review of the quality of patient information on the internet regarding inflammatory bowel disease treatments. Clin Gastroenterol Hepatol 2010;**8**(4):322-8.
14. Minervation. Minervation validation instrument for health care websites. http://lidaminervationcom/ 2011.
15. British Library and the University of Oxford. The DISCERN Instrument http://wwwdiscernorguk/discern_instrumentphp 2007(22-12-2014).
16. Charnock D, Shepperd S, Needham G, et al. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999;**53**(2):105-11.
17. Flesch R. A new readability yardstick. J Appl Psychol 1948;**32**(3):221-33.

18. Weis BD. *Health Literacy and Patient Safety: Help Patients Understand: Manual for Clinicians*: American Medical Association Foundation and American Medical Association, 2007.

19. Hargrave DR, Hargrave UA, Bouffet E. Quality of health information on the Internet in pediatric neuro-oncology. Neuro Oncol 2006;**8**(2):175-82.

20. Lieberman DA, Rex DK, Winawer SJ, 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. Sint Nicolaas J, de Jonge V, Cahen DL, et al. Awareness of surveillance recommendations among patients with colorectal adenomas. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2012;**10**(4):405-11.

22. Risoldi Cochrane Z, Gregory P, Wilson A. Readability of consumer health information on the internet: a comparison of U.S. government-funded and commercially funded websites. J Health Commun 2012;**17**(9):1003-10.

23. Grewal P, Alagaratnam S. The quality and readability of colorectal cancer information on the internet. Int J Surg 2013;**11**(5):410-3.

24. Hutchinson N, Baird GL, Garg M. Examining the Reading Level of Internet Medical Information for Common Internal Medicine Diagnoses. Am J Med 2016;**129**(6):637-9.

25. van Dam L, Kuipers EJ, Steyerberg EW, et al. The price of autonomy: should we offer individuals a choice of colorectal cancer screening strategies? Lancet Oncol 2013;**14**(1):e38-46.

26. Hol L, de Bekker-Grob EW, van Dam L, et al. Preferences for colorectal cancer screening strategies: a discrete choice experiment. Br J Cancer 2010;**102**(6):972-80.

27. Biesecker BB, Schwartz MD, Marteau TM. Enhancing informed choice to undergo health screening: a systematic review. Am J Health Behav 2013;**37**(3):351-9.

28. Wasserman M, Baxter NN, Rosen B, et al. Systematic review of internet patient information on colorectal cancer surgery. Dis Colon Rectum 2014;**57**(1):64-9.

29. Al-Bahrani A, Plusa S. The quality of patient-orientated internet information on colorectal cancer. Colorectal Dis 2004;**6**(5):323-6.

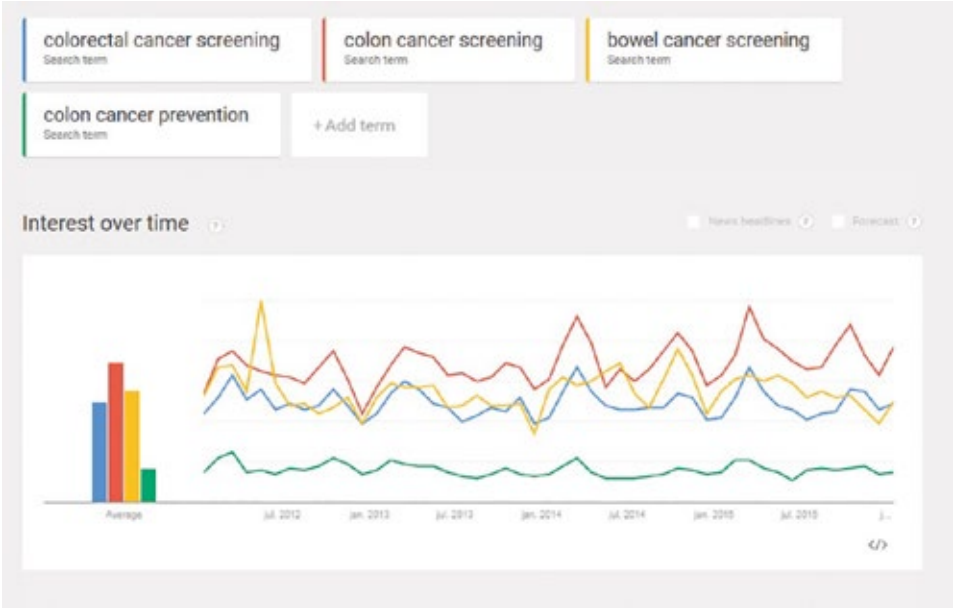
30. Saini SD. Website Review: Review of Patient-Oriented Websites for Colorectal Cancer Screening. Gastroenterology 2015;**148**(3):661-62.

31. Tian C, Champlin S, Mackert M, et al. Readability, suitability, and health content assessment of web-based patient education materials on colorectal cancer screening. Gastrointest Endosc 2014;**80**(2):284-90.

32. Chen CC, Yamada T, Smith J. An evaluation of healthcare information on the Internet: the case of colorectal cancer prevention. Int J Environ Res Public Health 2014;**11**(1):1058-75.

SUPPLEMENTARY FILES

Supplementary file 3.1 Google Trends over time for different search terms



Supplementary file 3.2 All scores per included website, sorted by their 2014 Google Rank.

Website URL	Google rank	Google rank	Google rank	WAS	Polyp Score	GQS	DISCERN	LIDA overall	LIDA accessibility	LIDA usability	LIDA reliability	FRE	FKG
	2014	2015	2016	points	points	points	points	(%)	(%)	(%)	(%)	school grade	points
www.cancerscreening.nhs.uk/bowel	1	9	12 *	24	5	3	45	75	87	75	53	10th	41
www.nhs.uk/Conditions/Cancer-of-the-colon-rectum-or-bowel/Pages/Screeningforbowelcancer.aspx	2	12	5	28	6	4	42	74	98	63	37	10th	55
www.cdc.gov/cancer/colorectal/basic_info/screening/	3	2	3	30	3	4	60	89	98	63	83	12th	35
www.nlm.nih.gov/medlineplus/ency/article/002071.htm	4	4	4	34	6	4	57	84	96	75	73	8th	59
www.bowelcanceraustralia.org/screening/	5	X	29	38	2	3	35	60	89	33	20	10th	58
www.cancer.org/cancer/colonandrectumcancer/moreinformation/colonandrectumcancerearlydetection/colorectal-cancer-early-detection-screening-tests-used	6	3	2	41	5	5	65	70	76	83	53	9th	62
www.cancercare.on.ca/colorectalscreeningχ	7	19	26	14	0	2	44	61	74	38	47	11th	40

Website URL	Google rank	Google rank	Google rank	WAS	Polyp Score	GQS	DISCERN	LIDA overall	LIDA accessibility	LIDA usability	LIDA reliability	FRE	FKG
	2014	2015	2016	points	points	points	points	(%)	(%)	(%)	(%)	school grade	points
www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/bowel-about	8	14	19	24	3	3	33	73	89	63	50	16th	27
www.cancer.gov/cancer-topics/pdq/screening/colorectal/Patient/page1	9	1	1	32	4	4	57	82	91	75	70	9th	62
www.mayoclinic.org/diseases-conditions/colon-cancer/in-depth/colon-cancer-screening/art-20046825	10	7	6	24	0	3	51	66	81	54	43	9th	53
http://healthfinder.gov/HealthTopics/Topic.aspx?id=15	11	10	10	12	1	3	30	74	85	54	63	5th	76
www.cancerresearchuk.org/cancer-info/spotcancerearly/screening/bowelcancerscreening/bowel-cancer-screening	12	11	13	25	0	2	41	76	96	58	47	10th	58
www.medicinenet.com/colon_cancer_screening/article.htm	17	8	9	34	10	4	58	77	87	67	63	10th	56
www.patient.co.uk/health/screening-for-colorectal-bowel-cancer	18	21 **	14 **	29	5	4	45	76	89	58	60	11th	48

Website URL	Google rank	Google rank	Google rank	WAS	Polyp Score	GQS	DISCERN	LIDA overall	LIDA accessibility	LIDA usability	LIDA reliability	FRE	FKG
	2014	2015	2016	points	points	points	points	(%)	(%)	(%)	(%)	school grade	points
www.macmillan.org.uk/Cancerinformation/Testsscreening/Bowelscreening/Bowelcancerscreening.aspx	19	X	X	35	5	4	49	72	93	67	37	11th	48
www.cancer.ca/en/prevention-and-screening/early-detection-and-screening/screening/screening-for-colorectal-cancer/?region=pe	21	X	X	17	0	3	45	64	87	46	30	12th	38
www.bccancer.bc.ca/PPI/Screening/colorectal.htm	23	X	X	10	0	2	43	26	0	67	57	13th	33
www.healthpei.ca/colorectal	25	X	X	9	1	2	29	66	85	42	40	11th	46
www.uptodate.com/contents/colon-and-rectal-cancer-screening-beyond-the-basics	27	13	16	37	6	5	69	88	83	88	97	14th	28
www.asge.org/patients/patients.aspx?id=8074	29	16	17	13	1	2	38	67	88	50	37	11th	42

X; Not in the first 3 Google pages results
* URL changed to https://www.gov.uk/topic/population-screening-programmes/bowel
** URL changed to http://patient.info/health/screening-for-bowel-colorectal-cancer

CHAPTER 7

Plenary feedback on inter-hospital differences to improve variation in quality of colonoscopy

E.H. Schreuders*
T.D.G. Belderbos*
M.A.C. Meijssen
R.J.T.H. Ouwendijk
T.J. Tang
F. ter Borg
P. van der Schaar
D.M. Le Fèvre
M. Stouten
E.A.M. Hassink
W.H. de Vos
P.C.J. ter Borg
M. Ledeboer
M.J. Bruno
L.M.G. Moons
E.J. Kuipers
P.D. Siersema
M.C.W. Spaander

*Both authors contributed equally to this manuscript.

Manuscript in preparation

ABSTRACT

Background

Quality of colonoscopy has been reported to vary considerably between institutions and endoscopists. Use of quality indicators can address suboptimal clinical outcomes and may reduce variation in colonoscopy performance. Therefore the aim of this study was to assess if plenary feedback on quality indicators can improve quality indicators of routine colonoscopy and reduce inter-hospital variation.

Methods

All consecutive routine colonoscopies performed in two academic and five non-academic hospitals in the Netherlands from November 2012 to January 2013 and November 2013 to February 2014 were prospectively registered. Between these two registration periods, a meeting with all seven hospitals was organized in which quality indicators and inter-hospital differences were discussed. Main quality indicators were adjusted cecal intubation rate (CIR) and adenoma detection rate (ADR). Quality indicators and variation in quality indicators were compared between hospitals in the first and second registration period.

Results

A total of 3129 patients were included in the first versus 5016 patients in the second registration period. The adjusted CIR was 97% in both registration periods. The overall ADR improved from 32% (95%CI: 30–34%) to 36% (95%CI: 34–37%) (P<0.001). Two out of the three hospitals had an ADR below average during the first registration period and improved to better-than-average. Higher inter-hospital variation was found for ADR (range in quality 25 – 47%; coefficient of variation (CV) 13.7) compared to adjusted CIR (range in quality 93 – 99%; CV 1.9). Variation in ADR reduced between the first and second registration period (CV 14.8 to 13.4).

Conclusions

In high-standard routine colonoscopy centers, plenary evaluation and discussion of inter-hospital differences reduced inter-hospital variation in quality of colonoscopy, although after correcting for confounders, overall ADR and adjusted CIR could not be improved.

BACKGROUND

Colonoscopy is the most common performed endoscopic procedure and considered the 'gold standard' investigation of the colon. The efficiency of colonoscopy in terms of reduction in post-colonoscopy colorectal cancer (CRC) incidence is mainly dependent on the quality of the procedure.[1 2] Quality of colonoscopy can be evaluated by measuring performance parameters of an individual endoscopist or group of endoscopists with predetermined targets. These predetermined targets are called quality indicators, which are measurable outcomes for which there is evidence-based impact on outcome that support a minimum standard.[3]

These quality indicators help to bring quality to higher standards and identifies suboptimal performance. Commonly used quality indicators are quality of the bowel preparation, cecal intubation rate, colonoscopy withdrawal time, and adenoma detection rate (ADR). The variation in ADR between endoscopists and the fact that ADRs is inversely correlated with the risk of post-colonoscopy colorectal cancers have become the rationale for the creation of targets for ADR in CRC screening.

Adjusted cecal intubation rate (CIR) and adenoma detection rate (ADR) are recognized by experts as priority quality indicators for colonoscopy.[3] For these indicators, there is a strong association between reaching the recommended target and important clinical outcomes.[1-3] The performance target is ≥25% for ADR and ≥90% for CIR.[3]

Variation in quality indicators between endoscopists plays an important role in colonoscopy outcome. Use of quality indicators and address suboptimal clinical outcomes, may reduce variations in colonoscopy performance, and will thereby close the gap between suboptimal and optimal performed colonoscopies. Therefore, reduction in variation of quality has emerged as an important priority for colonoscopy practice.[3] Data on inter-hospital variation in quality of colonoscopy performance and initiatives to reduce this variation are limited.[4-6] By measuring the quality indicators CIR and ADR differences in colonoscopy quality between hospitals can be depicted, as both indicators are affected by institutional and procedural factors.[7] Therefore, we set up a multicenter initiative comparing quality of routine colonoscopy in seven hospitals to incite quality improvement by plenary feedback and discussion on inter-hospital differences. The aim of this study was to assess whether plenary feedback on quality indicators can stimulate improvement of colonoscopy and thereby reduce inter-hospital differences.

METHODS

Study design and population

Design of the study and results of the baseline registration has been described elsewhere.[7] In short, we performed a prospective baseline registration (i.e. registration period 1), including all consecutive colonoscopies performed in two academic and five non-academic hospitals in the Netherlands between November 1st 2012 and January 10th 2013. All endoscopists filled out data on the indication for colonoscopy, type and dose of sedation used, quality of bowel preparation, cecal intubation, detection and removal of polyps, pathology reports and complications. After the first registration period, all quality indicators per hospital were evaluated and inter-hospital differences were discussed during a plenary session. No directives were given regarding quality improvement. A second prospective colonoscopy registration was conducted between November 1st, 2013 and February 24th, 2014.

Patient and colonoscopy characteristics were obtained from electronic medical records. Cecal landmarks were noted in the colonoscopy report, and photographic documentation was routinely obtained. Sedation was provided using midazolam, propofol, and/or opioid analgesics. We excluded patients with inflammatory bowel disease, polyposis syndromes or hereditary colorectal cancer syndromes and total colectomy. This research used a de-identified dataset and the study was exempt from obtaining informed consent from patients, as determined by the Medical Ethical Committee in accordance with the Medical Research Involving Human Subjects Act.

Colonoscopy procedure

Indications for colonoscopy were grouped into four categories: (1) anemia, abdominal symptoms and overt or occult rectal blood loss; (2) screening colonoscopies in patients with a positive family history of CRC; (4) surveillance after CRC or colorectal adenoma(s); and (5) other. The 'other' indication category consisted mostly of patients with diverticular disease, liver metastases or other abnormalities found during imaging. Complications were subdivided into bleeding (only taken into account if it did not stop spontaneously or by an intervention during the colonoscopy), perforation, and post-polypectomy syndrome. Endoscopists were gastroenterologists, fellows in training for gastroenterology, or nurse endoscopists. All endoscopists performing colonoscopies during the registration periods were informed about the study and consented to participate.

Definitions of quality indicators

Primary outcomes were the adjusted cecal intubation rate (CIR) and the adenoma detection rate (ADR). Secondary outcomes were adequate bowel preparation, mean number of adenomas per procedure (MAP) and mean number of adenomas per procedure in which at least one adenoma was detected (MAP+). The unadjusted CIR was defined as the proportion of colonoscopies in which the cecum was visualized, irrespective of reasons for not intubating the cecum. The adjusted CIR was calculated by excluding colonoscopies in which the endoscopist made the decision not to intubate the cecum because of severe colitis, colonic obstruction, or therapeutic targets not necessitating cecal intubation.[3] For the adjusted CIR, if no cecum was present (i.e. because of right-sided hemicolectomy or ileocecal resection) but the complete colon was visualized, the procedure was considered complete. Colonoscopies with poor bowel preparation were included in the adjusted cecal intubation rate, as preparation is considered to be part of colonoscopy practice. The ADR was defined as the proportion of procedures in which one or more adenomas were found. Adequate bowel preparation was defined as a Boston Bowel Preparation Scale (BBPS) score ≥6.[8] Split-dose bowel preparation was common practice in all seven hospitals.

Evaluation of variation

Variation in colonoscopy performance was evaluated between hospitals and between registration periods. Variations of performance parameters were analyzed quantitatively using coefficients of variation and visually in a funnel plot and colonoscopy quality indicator (CQI) plot. The funnel plots are a graphical representation of quality measures and provide additional insight in inter-hospital variation.[9 10] In the funnel plot, 95% confidence interval to the mean illustrate how much common cause variation is expected for a given hospital volume. After the first registration, the CQI plot was constructed to evaluate inter-hospital variation in CQI's and compare these with predetermined targets. The CIR and ADR per hospital and registration period were plotted in the CQI. For CIR, a minimum of 90% was used, as recommended by guidelines.[3] For ADR a line was drawn at the overall average of the first registration period (32%) to create a visual target for better-than-average performance.

Statistical analysis

Categorical variables with two categories were compared using the chi-squared test and those with multiple categories with binary logistic regression analyses. Differences between groups for continuous variables were tested using the t test.

We performed multivariate binary logistic regression analyses to assess whether the registration period was independently associated with adjusted CIR and ADR.

These analyses were adjusted for patient characteristics (gender, age, American Society of Anesthesiologists Physical Status Classification System (ASA) score) and procedure-related factors (indication, type of endoscopist, hospital, bowel preparation). Registration period and variables of potential significance ($P \leq 0.10$) were entered in the model.

Variation in quality indicators was quantified using coefficients of variation. We used the coefficient of variation because it allows the dispersion of variables with different means to be compared. The coefficient of variation in colonoscopy quality was measured as SD/mean, multiplied by 100.[11]

All p-values were two-sided and considered to be statistically significant if $P<0.05$. SPSS statistics software version 22.0 (IBM Corp., Armonk, New York, USA) was used for the calculations.

RESULTS

Baseline characteristics

A total of 3129 patients were included in the baseline registration (46% male, mean age 59 years) versus 5016 patients (49% male, mean age 60 years) in the second registration. Baseline characteristics are shown in Table 7.1. Median number of colonoscopies per hospital in the first registration was 568 (range 124- 793) and 670 (402 and 1375) in the second registration period. Most patients (70% and 65%) underwent colonoscopy because of anemia, abdominal symptoms or rectal (occult) blood loss. The majority of colonoscopies were performed by gastroenterologists with a median number of colonoscopies per endoscopist of 65 (IQR 41) in the first registration period and 103 (IQR 91) in the second registration period. Distribution of sex, age, ASA score and indication for colonoscopy differed significantly between the first and second registration period (Table 7.1).

Colonoscopy performance before and after plenary feedback meeting

All hospitals performed above the recommended target for adjusted CIR. The unadjusted and adjusted CIR did not change after plenary feedback and remained 95% ($P=0.81$) and 97% ($P=0.36$), respectively. Reasons for not intubating the cecum were inadequate bowel preparation (27%), stenosis or obstruction due to tumor or diverticulosis (29%), technical difficulty (27%), therapeutic goal not requiring cecal intubation (4.3%), pain (2.4%), ileocecal resection or right hemicolectomy (5.2%) or other (5.1%). The percentage of

patients with a BBPS ≥6 was 90% and increased to 91% (P=0.10). Percentage of polyp removal increased from 45% to 50% after plenary feedback (P<0.001) (Table 7.2). Two out of the three hospitals that had an ADR below average during the first registration period improved to better-than-average. Overall ADR improved from 32% to 36% (P<0.001). The MAP improved from 0.60±1.22 to 0.74±1.49 (P<0.001) and the MAP+ improved from 1.89 ± 1.48 to 2.09 ± 1.82 (P=0.005).

Table 7.1 Baseline characteristics

Registration period	1	2	P-value
Number of colonoscopies	3129	5016	
Age, mean ± SD (years)	59 ± 15	60 ± 15	0.006
Male gender	1423 (45.5)	2465 (49.2)	0.001
Indication (%)			
Symptomatic [a] / rectal (occult) blood loss	2198 (70.2)	3258 (65.0)	P<0.001
Family history of CRC	263 (8.4)	330 (6.6)	P=0.002
Surveillance after CRC / adenoma	626 (20.0)	1071 (21.4)	P=0.153
Other [b]	42 (1.3)	258 (5.1)	P<0.001
Unknown		96 (1.9)	
ASA-score (%)			
1	1784 (57.0)	2359 (47.0)	P<0.001
2	1061 (33.9)	2174 (43.3)	P<0.001
3	160 (5.1)	338 (6.7)	P=0.003
4	6 (0.2)	17 (0.3)	P=0.285
Unknown	118 (3.8)	128 (2.6)	P=0.002
Conscious sedation (%)			
Yes	2840 (90.8)	3783 (75.4)	P<0.001
No	158 (5.0)	110 (2.2)	
Not reported	131 (4.2)	1123 (22.4)	
Endoscopist (%)			
Gastroenterologist	1870 (59.8)	2804 (55.9)	P<0.015
Gastroenterology fellow	747 (23.9)	949 (18.9)	P<0.001
Nurse endoscopist	512 (16.4)	1164 (23.2)	P<0.001
Unknown		99 (2.0)	
Organization (%)			
1	339 (10.8)	618 (12.3)	
2	245 (7.8)	421 (8.4)	
3	124 (4.0)	402 (8.0)	
4	639 (20.4)	902 (18.0)	
5	568 (18.2)	670 (13.4)	
6	421 (13.5)	628 (12.5)	
7	793 (25.3)	1375 (27.4)	

[a] Includes anaemia, abdominal symptoms and changes in bowel habit

[b] Includes investigation for diverticular disease, abnormalities found with other imaging modalities and liver metastases.

American Society of Anesthesiologists Physical Status Classification System score (ASA score); Colorectal cancer (CRC); Standard deviation (SD)

Table 7.2 Variation in quality indicators for colonoscopy before and after plenary feedback.

Indicators	Before plenary feedback			After plenary feedback			
		SD	Coefficient of variation		p-value *	SD	Coefficient of variation
BBPS, median (IQR)	9 (3)			9 (2)	P=0.28		
BBPS ≥6, %(95% CI)	89.7% (88.5 – 90.7) (2697/3007)	3.0	3.3	90.8% (89.9 – 91.6%) (4543/5003)	P=0.10	2.9	3.2
Unadjusted CIR, (95% CI)	94.8% (94.0 – 95.5) (2967/3129)	2.2	2.3	94.3% (93.6 – 95.0) (4728/5013)	P=0.33	2.3	2.4
Adjusted CIR, (95% CI)	97.0% (96.4 – 97.6) (2967/3058)	1.7	1.8	96.7% (96.1 – 97.1) (4762/4927)	P=0.36	1.8	1.9
PDR, % (n)	45.4 % (1413)	5.0	11.0	50.2% (2518)	P<0.001	2.0	10.0
ADR, % (n)	31.8% (996)	4.7	14.8	35.7% (1792)	P<0.001	4.8	13.4
MAP, mean ± SD	0.60 ± 1.22	1.2	200	0.74 ± 1.49	P<0.001	1.5	201
MAP+, mean ± SD	1.89 ± 1.48	1.5	79.4	2.08 ± 1.84	P=0.005	1.8	86.5
Complications, % (n)	0.61% (19)			0.46% (23)	P=0.35		
Bleeding	0.29% (9)			0.34% (17)			
Post-polypectomy syndrome	0.06% (2)			–			
Perforation	0.03% (1)			0.06% (3)			
Other/unknown	0.22% (7)			0.06% (3)			

* P-value compared to quality indicator before feedback.
Interquartile range (IQR); 95% Confidence interval (95% CI); Cecal intubation rate (CIR); Polyp detection rate (PDR); Adenoma detection rate (ADR); Mean adenomas per procedure (MAP); Mean adenomas per positive procedure (MAP+); Standard deviation (SD)

Table 7.3 Outcomes per hospital and type of endoscopist

	Before and after plenary feedback	BBPS ≥6	CIR	Adj. CIR	ADR	MAP	Compli-cation
Hospital	**(n)**	**%**	**%**	**%**	**%**	**± SD**	**%**
1	1 (339)	79.0	96.2	98.5	27.4	0.50 ±1.00	1.2
	2 (618)	89.6	94.5	97.0	35.0	0.69 ±1.39	0.8
2	1 (245)	87.0	93.1	95.4	33.5	0.78 ±1.64	1.9
	2 (421)	86.5	94.8	96.4	40.9	0.99 ±1.74	1.0
3	1 (124)	88.7	99.2	99.2	46.8	1.00 ±1.44	0.0
	2 (402)	67.8	89.5	93.6	33.3	0.93 ±2.11	0.2
4	1 (639)	97.3	97.5	98.7	31.3	0.58 ±1.22	0.8
	2 (902)	95.2	97.2	98.2	34.1	0.68 ±1.32	0.3
5	1 (568)	95.8	98.1	99.3	24.8	0.42 ±0.97	0.2
	2 (670)	94.6	95.2	96.7	28.5	0.52 ±1.09	0.0
6	1 (421)	97.6	95.2	97.1	35.4	0.61 ±1.16	0.0
	2 (628)	95.5	95.5	97.9	42.4	0.88 ±1.52	0.6
7	1 (793)	80.2	89.4	93.4	34.4	0.68 ±1.27	0.6
	2 (1375)	92.4	92.6	95.9	36.7	0.74 ±1.46	0.4
Type of endoscopist							
Gastro-enterologist	1 (1870)	90.1	94.8	96.8	29.8	0.55 ±1.15	0.5
	2 (2804)	91.3	94.4	96.5	34.7	0.74 ±1.50	0.5
Fellow	1 (747)	85.7	94.2	97.0	34.0	0.75 ±1.49	1.2
	2 (949)	91.6	94.3	96.8	38.6	0.82 ±1.60	0.4
Nurse endoscopist	1 (512)	93.9	95.7	97.8	36.1	0.57 ±0.96	0.2
	2 (1164)	96.2	96.1	98.1	35.9	0.69 ±1.30	0.4

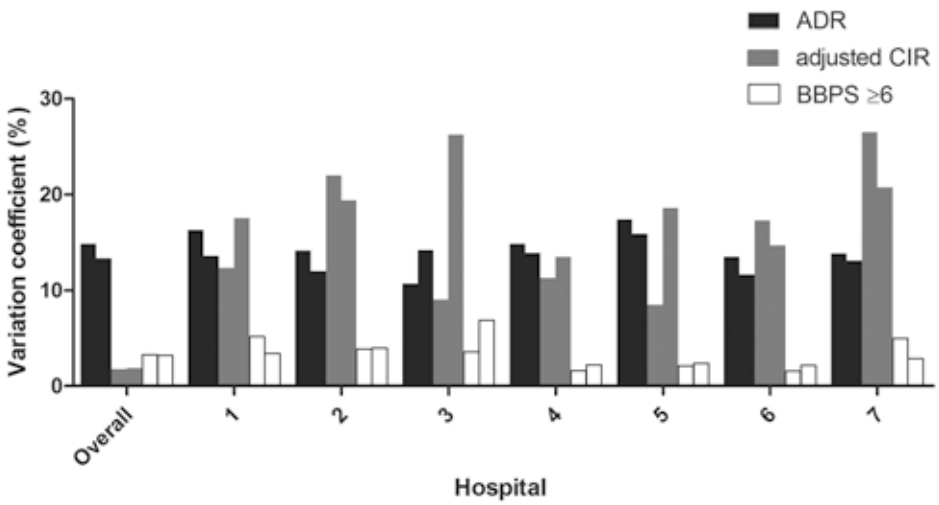
Before plenary feedback (1); After plenary feedback (2); Boston Bowel Preparation Scale (BBPS); CIR, cecal intubation rate; Adj. CIR, adjusted cecal intubation rate; ADR: adenoma detection rate.

Predictors of colonoscopy quality outcome

Table 7.4 shows the results of the multivariate logistic regression analyses. There was variation between registration periods in the odds of adjusted CIR. Adequate bowel preparation was strongly associated with the adjusted cecal intubation rate with an odds ratio (OR) of 33.65, $P<0.001$ (Table 7.4). ASA classification was inversely correlated with cecal intubation rate, in patients with an ASA score >1 the cecum was less often intubated. Hospital was an independent predictor for adjusted CIR ($P<0.001$) with OR's ranging from 1.22 (95%CI 0.73 – 2.05) to 3.86 (95%CI 2.15 – 6.93), with the hospital with largest volume as reference.

After correcting for confounders, performing colonoscopy after plenary feedback was not statistically significantly associated with ADR (OR 1.05, 95%CI 0.94 χ 1.17). In multivariate analysis, adenomas were more common in colonoscopies with an adequate bowel preparation (OR 1.73, 95%CI 1.38 – 2.17) and if cecal intubation was succeeded (OR 2.90, 95%CI 1.93 χ4.35) (Table 7.4). Highest ADR was found in colonoscopies performed for surveillance of CRC or adenomas, followed by a positive family history.

Figure 7.1 Variation of quality indicators among the 7 hospitals.



For each quality indicator, the left bar represents the first registration period and the right bar the second.

Table 7.4 Factors associated with adjusted cecum intubation rate and adenoma detection rate

	adjusted CIR		**ADR**	
	Adjusted OR (95% C.I.)	P	Adjusted OR (95% C.I.)	P
Registration period (first as reference)	0.70 (0.51 χ0.96)	0.03	1.05 (0.94 χ1.17)	0.39
Hospital		<0.001		<0.001
1	3.86 (2.15 χ6.93)		1.00 (0.82 χ1.21)	
2	1.22 (0.73 χ2.05)		1.18 (0.96 χ1.45)	
3	3.93 (1.76 χ8.76)		1.19 (0.93 χ1.53)	
4	1.79 (1.08 χ2.95)		0.83 (0.72 χ0.97)	
5	1.37 (0.81 χ2.32)		0.60 (0.50 χ0.73)	
6	1.35 (0.78 χ2.32)		1.08 (0.90 χ1.28)	
7	Reference		Reference	
Type of endoscopist		0.25		0.07
Gastroenterologist	Reference		Reference	
Fellow	1.26 (0.85 χ1.86)		1.15 (0.99 χ1.33)	
Nurse endoscopist	1.41 (0.89 χ2.24)		1.15 (1.00 χ1.33)	
Male sex			1.63 (1.47 χ1.81)	<0.001
Age, per 10-year increase	0.94 (0.84 χ1.05)	0.29	1.54 (1.48 χ1.61)	<0.001
ASA score		0.008		0.95
ASA 1	Reference		Reference	
ASA 2	0.59 (0.41 χ0.85)		1.01 (0.90 χ1.13)	
ASA 3	0.63 (0.37 χ1.06)		0.96 (0.77 χ1.20)	
ASA 4	0.16 (0.03 χ0.77)		1.22 (0.42 χ3.58)	
Indication		0.14		<0.001
Symptomatic / rectal (occult) blood loss	Reference		Reference	
Family history of CRC	1.29 (0.59 χ2.80)		1.45 (1.19 χ1.77)	
Surveillance after CRC / adenoma	1.61 (1.08 χ2.42)		1.78 (1.58 χ2.01)	
Other	1.08 (0.51 χ2.31)		1.44 (1.10 χ1.89)	
Adequate BBPS	33.65 (24.33 χ46.54)	<0.001	1.73 (1.38 χ2.17)	<0.001
Adjusted CIR	-		2.90 (1.93 χ4.35)	<0.001

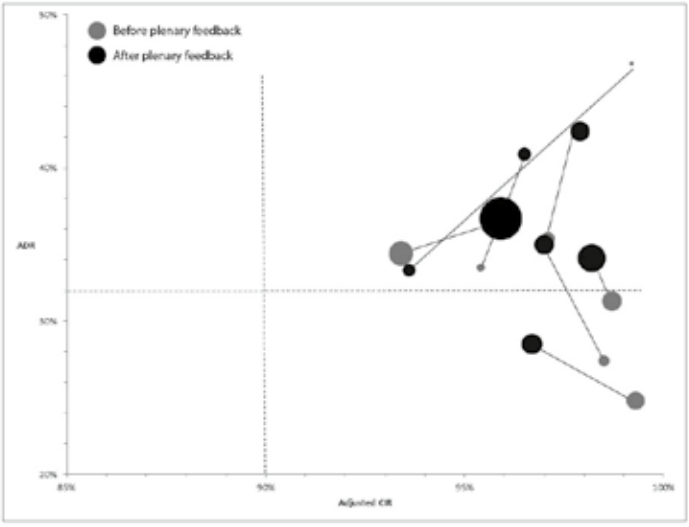
Odds ratio (OR); Confidence interval (CI); Cecal intubation rate (CIR); American Society of Anesthesiologists Physical Status Classification System score (ASA score); colorectal carcinoma (CRC); Boston Bowel Preparation Scale (BBPS).

Variation of quality measures among hospitals

The magnitude and pattern of variation in quality varied widely across the different quality indicators. Variations in all quality indicators between hospitals are depicted Table 7.3. In unadjusted analyses, there was significant inter-hospital variation in adjusted CIR ranging from 93% to 99% (P<0.001) before the plenary feedback and from 94% to 98% after the plenary feedback (P=0.001) (Table 7.3). Adequate bowel preparation varied between 79% and 98% (P<0.001) before plenary feedback and from 68% to 96% after plenary feedback (P<0.001). Inter-hospital ADRs varied between 25% and 47% (P<0.001) and between 29% and 41% (P<0.001) before and after plenary feedback respectively. Quality indicators with the smallest variation coefficient were adequate bowel preparation and unadjusted and adjusted CIR (Table 7.2). Overall variation coefficients decreased for adequate bowel preparation, PDR and ADR (Table 7.2).

Figure 7.1, showing the variation coefficients according to registration period and hospital, demonstrates large variation in the adjusted CIR between hospitals. Intra-hospital variation coefficients decreased for ADR in six out of the seven hospitals (Figure 7.1). The quality indicators with the smallest overall variation coefficient were adequate BBPS, and unadjusted and adjusted CIR (Table 7.2). Figure 7.2 shows in a CQI plot, inter-hospital variation in ADR and adjusted CIR. In the funnel plots, the mean ADR is depicted for each endoscopist (Figure 7.3) and hospital (Figure 7.4) by registration period, showing large variation between individual endoscopists and hospitals.

Figure 7.2 Colonoscopy quality indicator plot of ADR and adjusted CIR.

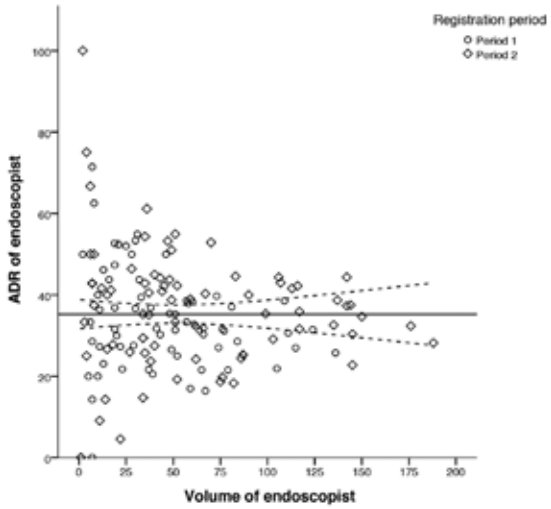


Each dot represents one hospital, bubble size is based on hospital-volume for that registration period

Adverse events

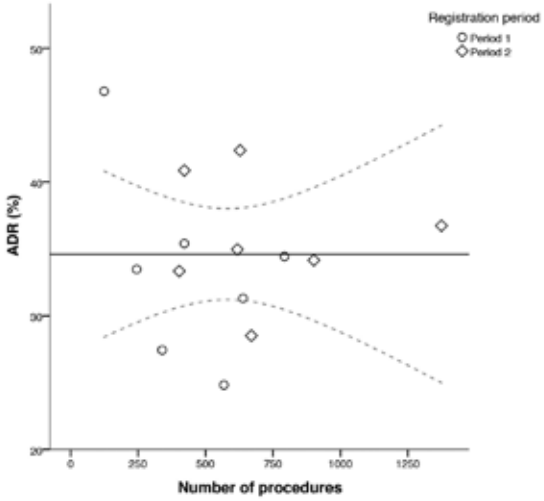
Plenary feedback did not affect complication rate (P=0.98). There were no procedure related deaths. The overall perforation rate was 0.06%.

Figure 7.3 Funnel plot showing each endoscopist's ADR.



Each dot represents an endoscopist in the first registration period, each rhombus an endoscopist in the second registration period. Horizontal line is the overall mean ADR (35%) with 95% confidence interval.
Adenoma detection rate (ADR).

Figure 7.4 Funnel plot showing each hospital's ADR.



Horizontal line is the overall mean ADR (35%) with 95% confidence interval.
ADR; Adenoma detection rate.

DISCUSSION

This multicenter initiative comparing quality of routine colonoscopy in seven hospitals revealed marked variation in quality indicators. We showed that a plenary feedback meeting to discuss inter-hospital differences, can reduce variation of quality between hospitals, although overall quality indicators could not be improved after correction for confounders. For hospitals with already good quality colonoscopies, plenary feedback helped to even improve colonoscopy performance. This underlines the importance of repeated quality assessment of colonoscopy in daily practice.

Implementation of CRC screening programs worldwide has largely increased colonoscopy demand.[12] This increases the relevance of quality assurance in colonoscopy. Evaluating quality indicators allows detecting differences in colonoscopy performance between hospitals and reducing variation. Variation in quality has now been convincingly linked to important outcome measures. For example, interval cancers are more common in endoscopists with low ADR as compared to those with high ADR.[1 2] In a large American study of 223,842 patients undergoing colonoscopy, a 3% reduction in CRC incidence and a 5% reduction in cancer mortality was found for each 1% increase in ADR.[2] Mainly based on this study, ADR targets have now been set at 25%, although guidelines specifically state this applies to asymptomatic, average risk individuals.[3] Age, gender, and indication of colonoscopy are known to affect the ADR, and adjustment for these factors is indicated when evaluating quality.[13 14] To improve ADR interventions targeting endoscopist performance have generally been ineffective.[4] Publically reporting physicians ADRs may improve ADR.[15] Although this study was done in a single private practice and did not correct for possible confounders. In our study we found that a performance target of ADR ≥25% appeared to be applicable in routine colonoscopies and an increase in ADR for a single hospital can be achieved after having a plenary feedback. It is however important to realize that these targets should not be considered standard of care for routine colonoscopies. Rather, they should be used as performance targets in the quality improvement process.[3]

Adjusted CIR was the strongest independent predictor of ADR in our study. This is also confirmed in previous studies showing that high CIR increases ADR, although contradictory conclusions have also been published.[16-18] In the second registration period only 68% of the procedures performed in hospital 3 had an adequate BBPS, whereas overall this percentage was 91%. This finding is mostly likely explained by the difference in case mix. The proportion ASA 3 or 4

patients was 26% in hospital 3 versus 6 % for the other hospitals together (data not shown). Inadequate bowel preparation in ASA class ≥3 has been described previously.[19 20] Also the strong decrease in performance in hospital 3 can be explained by case mix. The notable high ADR in the first registration period has been described previously.[7]

As expected and consistent with reports from other institutions, there was a substantial variation in ADRs between hospitals as well as between individual endoscopists.[7 14] After the plenary meeting, overall variation coefficients for ADR decreased and inter-hospital variation coefficients for ADR decreased in six out of the seven hospitals.

Improvement in performance is important but it is even more important to understand the factors responsible for improved performance. By understanding the factors that could be attributable to random (common-cause) variation in colonoscopy performance rather than to systematic variation in colonoscopy performance (which is an indication of a change in the system), opportunities for targeted educational programs can be identified. Explanations that possible attributed to the improvement found in our study could be increased awareness and feedback. There is substantial evidence that feedback can effectively improve quality.[21 22] However, it is known that feedback may be more effective when baseline performance is low, which does not apply for this study. Another reason for the improvement may be the start of the national Dutch colorectal cancer screening program in 2014, for which strict certification and continuous monitoring of quality measures of endoscopists and hospitals is required.

To appreciate these findings, some limitations need to be discussed. Because endoscopists were aware of the quality registration, a Hawthorne effect may be contributing to the observed result.[23] Another limitation of the study is that the sample size of colonoscopies from each endoscopist (median 65, IQR 41) may be insufficient to reliably assess individual ADR. We were unable to determine whether inter-hospital differences with regard to interventions attributed to improvement.

Different studies examined the effect of different interventions on the performance of colonoscopy quality.[4-6] Our findings suggest that endoscopists' knowledge that quality measures are being monitored represents a powerful tool that by itself can stimulate improvement of quality of colonoscopy. There the MAP and MAP+ reflect inspection of the entire length of the colon better than the ADR which rewards a "one-and-done" approach to colonoscopy, it is suggested in literature that it provides greater differentiation between endoscopists.[24] This is

also in line with our results showing that both MAP and MAP+ had the highest variation coefficient and inter-hospital variance. Our findings encourage using MAP and MAP+ to assess quality variation in colonoscopy performance. Besides the use of MAP and MAP+ future studies should focus on qualitative evaluation of feedback pathways that have led to improvement. Understanding these pathways may allow the same or similar strategies to be applied in different settings. Furthermore, it is important to consider a longer-term view of performance and assess changes over a long time period. Strategies theretofore that account for regression-to-the-mean have been suggested.[25]

In conclusion, there is considerable variation in quality indicators for colonoscopy. However, targets for CQI in CRC screening programs can also be achieved in routine colonoscopies. By plenary evaluation and discussion of inter-hospital differences in colonoscopy performance, variation between hospitals, especially ADR, can be improved. Further research is needed to identify factors responsible for the variation between hospitals, to achieve optimal effectiveness in colonoscopy performance.

REFERENCES

1. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;**362**(19):1795-803.
2. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;**370**(14):1298-306.
3. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. Am J Gastroenterol 2015;**110**(1):72-90.
4. 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.
5. Shaukat A, Oancea C, Bond JH, et al. Variation in detection of adenomas and polyps by colonoscopy and change over time with a performance improvement program. Clin Gastroenterol Hepatol 2009;**7**(12):1335-40.
6. Abou Fadel CG, Shayto RH, Sharara AI. Optimizing Colonoscopy Quality: From Bowel Preparation to Surveillance. Curr Treat Options Gastroenterol 2016;**14**(1):115-27.
7. Belderbos TD, Grobbee EJ, van Oijen MG, et al. Comparison of cecal intubation and adenoma detection between hospitals can provide incentives to improve quality of colonoscopy. Endoscopy 2015;**47**(8):703-9.
8. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc 2010;**72**(4):686-92.
9. Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med 2005;**24**(8):1185-202.
10. Neuburger J, Cromwell DA, Hutchings A, et al. Funnel plots for comparing provider performance based on patient-reported outcome measures. BMJ Qual Saf 2011;**20**(12):1020-6.
11. le Clercq CM, Mooi RJ, Winkens B, et al. Temporal trends and variability of colonoscopy performance in a gastroenterology practice. Endoscopy 2016.
12. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.
13. Ferlitsch M, Reinhart K, Pramhas S, et al. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA 2011;**306**(12):1352-8.
14. Corley DA, Jensen CD, Marks AR, et al. Variation of adenoma prevalence by age, sex, race, and colon location in a large population: implications for screening and quality programs. Clin Gastroenterol Hepatol 2013;**11**(2):172-80.
15. Abdul-Baki H, Schoen RE, Dean K, et al. Public reporting of colonoscopy quality is associated with an increase in endoscopist adenoma detection rate. Gastrointest Endosc 2015;**82**(4):676-82.
16. Zorzi M, Senore C, Da Re F, et al. Quality of colonoscopy in an organised colorectal cancer screening programme with immunochemical faecal occult blood test: the EQuIPE study (Evaluating Quality Indicators of the Performance of Endoscopy). Gut 2015;**64**(9):1389-96.
17. Jover R, Zapater P, Polania E, et al. Modifiable endoscopic factors that influence the adenoma detection rate in colorectal cancer screening colonoscopies. Gastrointest Endosc 2013;**77**(3):381-89 e1.
18. de Jonge V, Sint Nicolaas J, Cahen DL, et al. Quality evaluation of colonoscopy reporting and colonoscopy performance in daily clinical practice. Gastrointest Endosc 2012;**75**(1):98-106.

19. Yadlapati R, Johnston ER, Gregory DL, et al. Predictors of Inadequate Inpatient Colonoscopy Preparation and Its Association with Hospital Length of Stay and Costs. Dig Dis Sci 2015;**60**(11):3482-90.

20. Dik VK, Moons LM, Huyuk M, et al. Predicting inadequate bowel preparation for colonoscopy in participants receiving split-dose bowel preparation: development and validation of a prediction score. Gastrointest Endosc 2015;**81**(3):665-72.

21. Rajasekhar PT, Rees CJ, Nixon C, et al. Factors influencing change in clinical practice: A qualitative evaluation of the implementation of the quality improvement in colonoscopy study. Int J Health Care Qual Assur 2016;**29**(1):5-15.

22. Ivers NM, Grimshaw JM, Jamtvedt G, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. J Gen Intern Med 2014;**29**(11):1534-41.

23. Fletcher R. *Clinical epidemiology: the essentials*. 3rd ed: Williams & Wilkins, 1996.

24. Barclay RL, Vicari JJ, Doughty AS, et al. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006;**355**(24):2533-41.

25. Kasza J, Moran JL, Solomon PJ. Assessing changes over time in healthcare provider performance: addressing regression to the mean over multiple time points. Biom J 2015;**57**(2):271-85.

CHAPTER 8

The appropriateness of surveillance colonoscopy intervals after polypectomy

E.H. Schreuders*
J. Sint Nicolaas*
V. de Jonge
H. van Kooten
I. Soo
D. Sadowski
C. Wong
M.E. van Leerdam
E.J. Kuipers
S.J.O. Veldhuyzen van Zanten

*Both authors contributed equally to this manuscript.

Canadian Journal of Gastroenterology. 2013 Jan;27(1):33-8.

ABSTRACT

Introduction

Adherence to surveillance colonoscopy guidelines is important to prevent colorectal cancer (CRC) and unnecessary workload. This study evaluated how well Canadian gastroenterologists adhere to colonoscopy surveillance guidelines after adenoma removal or treatment for colorectal cancer.

Methods

Patients with a history of adenomas or CRC who had surveillance performed between October 2008 and October 2010 were retrospectively included. Time-intervals between index colonoscopy and surveillance were compared to the guidelines of the American Gastroenterology Association (AGA) of 2008 and regarded as appropriate when the surveillance interval was ±6 months of the recommended time interval.

Results

265 patients were included (52% male; mean age 58 years). Among patients with a normal index colonoscopy (n=110), 42% received surveillance on time, 38% too early (median difference=1.2 years too early), and 20% too late (median difference=1.0 years too late). Among patients with non-advanced adenomas at index (n=96), 25% received surveillance on time, 61% too early (median difference=1.85), and 14% too late (median difference=1.1). Among patients with advanced neoplasia at index (n=59), 29% received surveillance on time, 34% too early (median difference=1.86), and 37% later than recommended (median difference=1.61). No significant difference in adenoma detection rates was observed when too early surveillance vs. appropriate surveillance (34 vs. 33%, p=0.92) and too late surveillance vs. appropriate surveillance (21% vs. 33%, p=0.11) were compared.

Conclusion

A minority of surveillance colonoscopies are being performed according to guideline recommendations. Deviation from the guidelines did not improve the adenoma detection rate. Interventions aimed at improving adherence to surveillance guidelines are needed.

INTRODUCTION

Colorectal cancer (CRC) is a leading cause of cancer-related mortality in the Western world.[1] Screening for CRC decreases CRC-related mortality and CRC incidence.[2] The adenoma-carcinoma sequence is accepted as the pathway of development of CRC and hence one of the main aims of screening colonoscopy is to detect and completely remove all adenomas.[3 4] After neoplasia removal, patients remain at increased risk for adenoma recurrence. Therefore, surveillance after removal of adenomas or CRC is recommended. Some factors are associated with an increased risk of adenoma recurrence such as number of previous polyps and presence of villous features on histology.[5] The surveillance interval is generally based on these findings at the index colonoscopy.[6]

The demand for colonoscopy procedures has risen considerably over the last years which led to increased wait times for gastroenterology care in many regions of the world, including Canada.[7-9] The increase in demand for colonoscopy as a part of CRC screening is likely to lengthen wait times even further. Deviation from surveillance guidelines may further lead to unnecessary workload, and consequently a decrease in cost-effectiveness of CRC screening.[10] Previous studies have shown that a significant proportion of gastroenterologists recommend follow-up intervals that deviate considerably from the published guidelines.[11-13]

The objective of this study was to assess the appropriateness of recommended surveillance colonoscopy intervals in the Canadian endoscopy setting.[6] Furthermore, we aimed to determine whether the appropriateness of surveillance intervals influenced the detection of colorectal adenomas.

METHODS

This retrospective cohort study was conducted at the University of Alberta Hospital, Edmonton, Alberta, Canada. Ethical approval for this study was obtained by the Health Research Ethics Board (Pro00013953). Patients were identified and selected from a pilot study carried out as a first step in the creation of a CRC screening program (NCT00893503). This screening program, called SCOPE (Stop Colorectal Cancer through Prevention and Education), was launched in Edmonton to start a regional colon cancer screening program. The program was designed to test several steps in the referral process. The average risk patient could be referred only if they had a positive fecal occult blood test. Patients were

also eligible to be referred to the program if they had a personal history of colon cancer or adenomatous polyps, or a family history of colon cancer or polyps. In the pilot study the program only accepted referrals coming from gastroenterologists. In all patients who had a personal history of cancer or adenomatous polyps, the baseline endoscopy report and histology of removed polyps was available. For all patients who had a prior index colonoscopy during which adenomatous polyps were removed, the program did accept the recommendation that was made by the colonoscopist who performed the index colonoscopy. Patients where only included in the current study if they had a personal history of adenomas or CRC and underwent colonoscopy in the SCOPE program for surveillance purposes. Patients with a history of inflammatory bowel disease, a known hereditary CRC syndrome or patients with colonoscopies that were performed for the evaluation of gastrointestinal symptoms were excluded.

Patients with a personal history of adenomas or CRC who had a surveillance colonoscopy performed between October 2008 and October 2010 were included. The colonoscopy performed before the procedure done between October 2008 and October 2010 was defined as the index colonoscopy. As all patients had an adenoma history this index colonoscopy might not have been their actual first-time colonoscopy done for adenoma or CRC surveillance. Consequently, even if our defined index colonoscopy was normal, these patients according to the American Gastroenterogical Association (AGA) guidelines of 2008 were supposed to undergo surveillance colonoscopy every 5 years because of their adenoma or CRC history.[6] This aspect was not incorporated in the 2006 AGA guideline but was already indebted in the 2006 American Society for Gastrointestinal Endoscopy (ASGE) guideline.[14] The 2008 AGA guidelines were used in our analyses as this was a combination of the AGA and ASGE guidelines from 2006, both available at that time.

Data collection

The following data were collected from endoscopy reports: demographic data (age and gender), family history for CRC, index and surveillance colonoscopy characteristics such as date, cecal intubation rate, quality of bowel preparation (if not mentioned in the report it was assumed to be sufficient), and endoscopic findings including diagnosis, number, histology and site of polyps or cancer. Right sided adenomas were defined as adenomas found in the cecum, ascending colon, hepatic flexure or transverse colon. Left sided was defined as splenic flexure, descending colon, sigmoid colon and rectum. Patients were categorized in different surveillance groups based on their most advanced lesion at index colonoscopy: normal, non-advanced adenoma, or advanced neoplasia. Advanced

neoplasia was defined as ≥3 adenomas or adenomas >10mm, with >25% villous histology or high-grade dysplasia, or CRC. For patients who were diagnosed with CRC during index colonoscopy, the date of their surgery was used in order to calculate the optimal surveillance interval.

The actual interval between the index and surveillance colonoscopy was compared to the recommended interval stated in the 2008 guidelines from the AGA.[6] This guideline was used as the Canadian guideline, which has not been updated since 2004, does not state explicit recommendations for surveillance. A margin of six months around the recommended date was considered as an appropriate surveillance interval. Outcome measures were defined as the percentage of appropriate, too early, and too late procedures. Secondary outcomes were the adenoma detection rates (ADR) of the three categories, defined as the proportion of patients who had at least one adenoma at surveillance colonoscopy. For the ADR analyses and appropriateness categories, it was deemed to be appropriate to exclude the cases with poor bowel preparation on index procedure.

Statistical analysis

Descriptive statistics were used. Differences were assessed for significance by means of the Student's t-test for continuous data and the Chi-square test for categorical data. The level of statistical significance was defined as a two-sided p-value <0.05. All analyses were performed using statistical software package SPSS PASW 17.0, Chicago, IL, USA.

RESULTS

After excluding 11 cases in whom no information was available about the index findings, 265 patients were included for analyses (52% male; mean age on index: 58 yrs, SD=11). Table 8.1 summarizes the patients' characteristics stratified for the findings at index colonoscopy. The median number of previous colonoscopies was 1 (range: 0-6). Index colonoscopy was normal in 42% of the patients (n=110/265), non-advanced adenomas were found in 36% (n=96/265), and advanced neoplasia was detected in 22% of the cases (n=59/265). Three patients (1%) had CRC.

Table 8.1 Patient characteristics at index colonoscopy

	Total (n=265)	Normal findings (n=110)	Nonadvanced adenoma (n=96)	Advanced neoplasia (n=59)
Age, years, mean ± SD	58±10.5	59±10.8	57±10.0	59±11.0
Male sex, n (%)	138 (52)	61 (56)	46 (48)	31 (53)
Cecal intubation rate, n (%)	230 (95)	98 (94)	84 (97)	48 (94)
Adequate bowel preparation, n (%)	238 (90)	97 (88)	89 (90)	55 (93)
Family history of CRC*, n (%)	61 (23)	22 (19)	27 (28)	13 (22)
Interval until surveillance colonoscopy (years, ±SD)	3.8±1.7	4.3±1.5	3.8±1.5	2.8±2.0

* first degree relatives with CRC extracted from endoscopy report if available

Surveillance colonoscopy

Of all 265 surveillance colonoscopies, 33% (n=87/265) were classified as procedures performed on time according to the AGA guidelines. In 46% of the patients (n=121/265) the surveillance interval was shorter than recommended, and the remaining 21% (n=57/265) underwent surveillance later than recommended compared to the surveillance guidelines.

Figure 8.1 shows the actual observed mean time interval between index and surveillance colonoscopy compared with the recommended time interval stratified for the index finding. The median difference between the recommended time interval and the observed interval was -1.8 years (inter quartile range (IQR)=1.12) for surveillance colonoscopies which were performed too early, and +1.1 years (IQR=1.34) for those which were performed too late.

Table 8.2 shows the findings at index colonoscopy per appropriateness category. In 17% of the patients who received surveillance colonoscopy too early, a poor bowel prep quality was mentioned at index procedure compared to 5% at surveillance procedures performed on time ($p<0.01$). No significant differences for a positive family history and cecal intubation rates were observed between the three appropriateness categories (all p-values > 0.1). After exclusion of patients with poor bowel preparation on index, the proportion of patients with a surveillance colonoscopy performed on time was 35%.

Figure 8.1 Mean time interval between index and surveillance colonoscopy compared to time interval recommended by the AGA guideline of 2008.

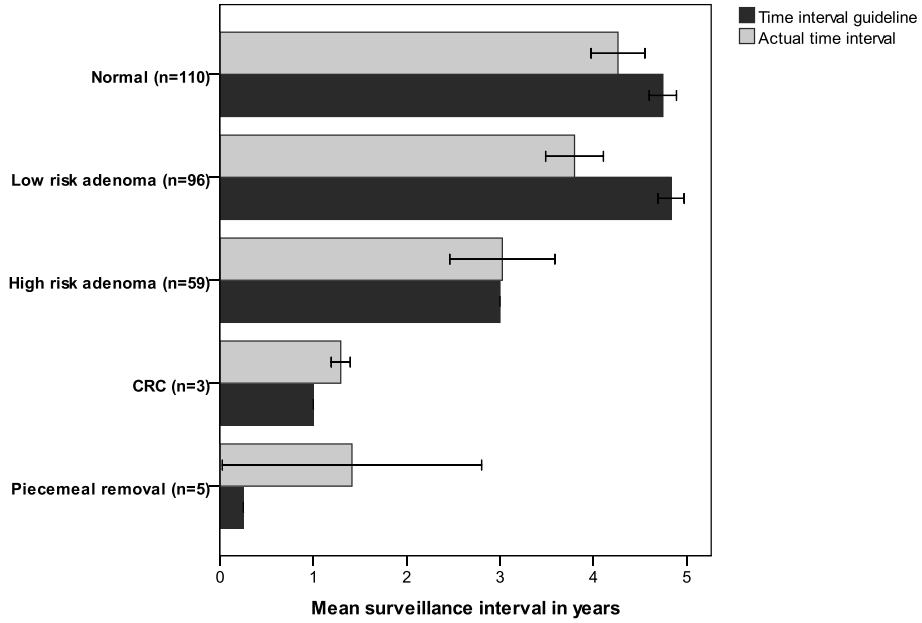


Table 8.2 Findings at index stratified for appropriateness according to the guidelines

	Surveillance colonoscopy, n (%)		
	On time (n=87)	**Too early* (n=121)**	**Too late* (n=57)**
Positive family history of CRC	17 (20)	29 (24)	15 (26)
Cecal intubation rate at index #	78 (93)	111 (96)	42 (98)
Adequate bowel prep at index	83 (95)	100 (83) **	55 (97)
Median difference, years (IQR)* **	0.07 (0.40)	-1.8 (1.12)	1.1 (1.34)
Findings at index			
Normal index	46 (53)	42 (35) **	22 (39)
Non advanced adenoma at index	24 (27)	59 (49) **	13 (22)
Advanced neoplasia∞ at index	17 (20)	20 (17)	22 (39) **

*= earlier respectively later than recommended in the 2008 AGA guideline; **= statistically significant compared with surveillance colonoscopy on time; ***=mean difference between the recommended time interval and the observed interval between index and surveillance colonoscopy; #= total numbers differ due to missing data; ∞= ≥3 adenomas or >10mm, with (tubulo-)villous histology or high-grade dysplasia or colorectal cancer. IQR inter quartile range; prep Preparation

Normal index colonoscopy

Among patients with a normal index colonoscopy (n=110), 42% of patients (n=46/110) received surveillance colonoscopy on time, 38% (n=42/110) too early (median difference= -1.23 years too early; IQR=1.54), and 20% (n=22/110) too late (median difference=0.98 too early; IQR=0.86).

Non advanced adenoma at index colonoscopy

Among patients with non-advanced adenomas at index colonoscopy (n=96), 25% (n=24/96) received surveillance colonoscopy on time, 61% of cases (n=59/96) too early (median difference= -1.85 years too early; IQR=1.24), and 14% (n=22/96) of patients received their surveillance colonoscopy too late (median difference=1.05 years too late; IQR=1.74). Among patients with normal or non-advanced findings on index colonoscopy (n=206), the percentage of patients undergoing surveillance earlier than recommended was significantly higher (49%; n=101/206) than in patients with advanced neoplasia (34%; n=20/59) (p=0.040).

Advanced neoplasia at index colonoscopy

Among patients with advanced neoplasia, 29% (n=17/59) received surveillance colonoscopy on time, 34% (n=20/59) too early (median difference= -1.86 years too early; IQR=0.92), and 37% (n=22/59) later than recommended (median difference=1.61 years too late; IQR=2.08). All patients that had CRC detected at index colonoscopy (n=3) did receive surveillance on time. Of the five patients who had a piecemeal removal of their advanced adenoma at index colonoscopy, four returned too late and two of these patients had advanced neoplasia at surveillance colonoscopy. Among patients with advanced neoplasia on index (n=59), the percentage of patients undergoing surveillance later than recommended was significantly higher (37%; n=22/59) than in patients with non-advanced or normal findings (17%; n=35/206) (p=0.001).

Adenoma location distribution on index procedure and surveillance practice patterns

In Table 8.3, the distribution in location of most advanced index findings is presented, stratified for the different appropriateness categories. Overall, in 53% the adenoma location was only right-sided. Among patients who were diagnosed with both left and right-sided adenomas at index, earlier surveillance interval than recommended by guidelines was more often seen, compared to surveillance on time according to the guidelines (8% vs. 2% p=0.08).

Table 8.3 Adenoma findings and sites on index and surveillance colonoscopy stratified for appropriateness according to the guidelines

	Surveillance colonoscopy, n (%)		
	On time (n=87)	**Too early* (n=121)**	**Too late* (n=57)**
Right sided adenomas index	24 (28)	38 (31)	19 (33)
Left sided adenomas index	17 (20)	29 (24)	14 (25)
Adenomas on both sides index	2 (2)	10 (8)	1 (2)
Right sided adenomas SC	22 (25)	22 (18)	9 (16)
Left sided adenomas SC	5 (6)	13 (11)	5 (9)
Adenomas on both sides SC	3 (3)	10 (8)	9 (0)

SC= surveillance colonoscopy; *= earlier respectively later than recommended in the 2008 AGA guideline;

Adenoma detection rate at surveillance colonoscopy

Adenomas were found on surveillance in 32% of cases (n=83/265), 8% of patients (n=20/265) had advanced adenomas. No CRC was identified at surveillance colonoscopy.The ADR for appropriate versus non-appropriate surveillance colonoscopy (stratified for index findings) are summarized in Table 8.4. Patients were excluded from this analysis, if they underwent surveillance procedure too early and had a poor bowel preparation during the index colonoscopy. The ADR at surveillance colonoscopy was significantly higher in patients with advanced neoplasia at index (n=26/59, 44%) vs. normal index colonoscopy (n=26/100, 26%), p=0.01. No significant difference in the ADR on surveillance was observed for procedures that were performed on time according to the guidelines compared to too early performed procedures (33% (n=29/87) vs. 34% (n=34/100 respectively), p=0.923). The ADR was also not significantly different between appropriate versus too late procedures (33% (n=29/87) vs. 21% (n=12/57) respectively, p=0.11). The detection of advanced adenomas at surveillance colonoscopy was not significantly different between appropriate vs. too early performed procedures (5% (n=4/87) vs. 10% (n=10/100), p=0.161) nor for appropriate vs. too late performed surveillance according to the guidelines (5% (n=4/87) vs. 9% (n=5/57), p=0.312).

Table 8.4 Adenoma detection rate at SC stratified for appropriateness according to the guidelines

	Surveillance colonoscopy, n (%)		
	On time (n=87)	**Too early* (n=121)**	**Too late* (n=57)**
ADR at SC			
Adenoma at SC	29 (33)	24 (34)	12 (21)
Nonadvanced adenoma at SC	25 (29)	24 (24)	7 (12)
Advanced neoplasia∞ at SC	4 (5)	10 (10)	5 (9)
ADR at SC per index			
Normal index	14 (16)	8 (8)	4 (7)
Non-advanced adenoma index	8 (9)	14 (14)	4 (7)
Advanced neoplasia at index	7 (8)	12 (12)	4 (7)

*= earlier respectively later than recommended in the 2008 AGA guideline; ∞= ≥3 adenomas or >10mm, with (tubulo-)villous histology or high-grade dysplasia or colorectal cancer; #= poor bowel preparation on index excluded

DISCUSSION

Recent reports have shown that there are significant problems with wait times for colonoscopy procedures in many centers in Canada.[7] It is expected that, in the context of CRC screening and its associated need for surveillance procedures, the demand on endoscopy units will increase. This study aimed to assess how well endoscopists in the Canadian endoscopy setting adhere to the guidelines for surveillance colonoscopies and whether improvements are achievable which would help to decrease wait times.

Our study showed that in a significant proportion of patients surveillance colonoscopy was not performed at the recommended time interval. Only 33% of the patients underwent a surveillance colonoscopy according to the AGA guideline. The largest group consisted of patients that underwent procedures earlier than recommended (46%). Underuse was also reported, as 21% of the patients received their colonoscopy too late. Shortening or lengthening the surveillance intervals did not significantly affect the ADR.

Several surveys have documented suboptimal usage of surveillance colonoscopy, with physicians often recommending surveillance intervals that are too short.[12 15] A Dutch study reported that 52% of the respondents used shorter surveillance intervals than stated by the national recommendations.[11] Suboptimal usage of adherence in daily practice has also been shown in several studies.[13 16-18] A study from the USA observed a considerable disparity between guidelines and endoscopists' recommendations in colonoscopy reports, with more surveillance colonoscopies occurring too soon; in only 37% of the cases were the recommendations consistent with the guidelines.[13] Another study from the Netherlands reported low follow-up rates for surveillance colonoscopy after the removal of adenomas or CRC; more than one-third of patients (35%) tended not to undergo surveillance colonoscopies although overuse was also observed.[17]

As the risk for finding adenomas during surveillance colonoscopy differs based on baseline findings, guideline recommendations for surveillance colonoscopy are stratified based on the index findings.[5 6] There is evidence that surveillance colonoscopy is over-utilized in low-risk subjects and underutilized in high-risk subjects.[16] A US community practice assessment of utilization of surveillance colonoscopy showed under-usage of surveillance practice in terms of longer follow up intervals if high risk lesions at index colonoscopy existed (31%).[13] In our study a similar trend was observed in adherence patterns for surveillance practice between advanced and non-advanced lesions on index procedures.[18] Patients with non-advanced adenomas (49%) often received surveillance too early while patients with advanced neoplasia often underwent surveillance colonoscopy too late according to the guidelines (37%).

Of all patients with index procedures that revealed advanced adenoma, 39% also had adenomas at surveillance colonoscopy (n=23/59). This underlines that advanced adenoma at index colonoscopy is an important risk factor for adenoma recurrence and thereby supports the guidelines for more vigilant surveillance.[16] However, in our study the detection rate of recurrent adenomas was also high in 25% of patients with normal index colonoscopies. It emphasizes that our current findings must be taken with caution for surveillance colonoscopy in general. The fact that 25% of the patients with a normal index still had adenomas at surveillance colonoscopy indicates that these patients remain at high risk for developing metachronous adenomas, despite normal findings at a previous surveillance colonoscopy. The yield of surveillance colonoscopy did not significantly differ between colonoscopies carried out at appropriate or inappropriate times, suggesting that deviating from the guidelines does not necessarily affect the yield of surveillance colonoscopy. However, our sample size may have been too

small to detect significant differences in ADR. In addition, similar detection rates in advanced adenomas between on time (8%) vs. too late procedures (7%) were observed. It is well established that the detection of adenomas is dependent on the quality of bowel preparation.[19] Clinical decisions about the surveillance interval derived from colon cleanliness assessment can vary considerably among endoscopists and there is little agreement on what constitutes an insufficient bowel preparation.[20] However, in our analyses it was shown the surveillance procedures performed too early still yielded high detection rates appreciably if patients with a poor bowel preparation were excluded.

Apart from suboptimal bowel preparation on index procedure, several other explanations have been suggested for the high detection rate and non-adherence to surveillance recommendations such as an incomplete examination, possibly incomplete removal of lesions and the presence of a family history of CRC.[15] Although in the too early surveillance cohort relatively more patients had a family history of CRC (24%) compared to the surveillance on time population (20%), the difference was not significant. It must be acknowledged that we do not know how reliable the reporting is for family history in colorectal cancer in this retrospective analysis. Additionally, there were no significant differences between cecal intubation rates in the three appropriateness categories. Quality issues may have been involved but these could not be adequately assessed in a prospective manner. In the cohort of patients that received earlier surveillance, 8% of patients had left- and right-sided adenomas versus 2% in the cohort that received surveillance on time ($p=0.08$). Retrospective reports have argued that the effectiveness of colonoscopy for left-sided and right-sided colorectal neoplasia seems to differ. A Canadian study showed that the protective effect of a complete colonoscopy was strong for mortality from distal lesions, but not associated with mortality from proximal lesions.[21] This might be an explanation why physicians recommended a shorter surveillance time interval if patients had lesions in the proximal part of their colon or to detect synchronous lesions in the proximal and distal colon. However, our data did not show that recommendations for only proximal-sided lesions were shorter compared to surveillance if only distal lesions on index colonoscopy existed.

Additionally, insufficient awareness of guidelines may be an important factor for non-adherence by physicians. Several studies have shown that appropriate use of surveillance after the detection of adenomas or CRC depends to a great extent on the knowledge physicians have of surveillance guidelines.[12 15] A recent study using hypothetical cases evaluating the knowledge of Canadian endoscopists about guidelines for follow-up colonoscopies showed that many gave the wrong

recommendation.[22] Another study showed that priming endoscopists by means of distributing guideline pocket pamphlets for use in endoscopy units, did increase the compliance to guidelines.[23]

Another possible explanation why endoscopists may recommend follow up colonoscopies that are too soon is that they base the recommendation on the number of polyps that they removed during the procedure before the pathology is back. An example of this would be a patient with four small polyps but the pathology only showed adenomas in two. Other explanations for less effective surveillance programs besides nod-adherence by physicians can be found in patient related factors such as non-attendance to surveillance colonoscopy. Most studies in this area focus on clinician adherence to published guidelines rather than patient adherence to clinician recommendations. Because our study design was limited to only patients who had returned for their surveillance colonoscopy, it is not known how many patients, who underwent an index colonoscopy that warranted follow-up, did not return for surveillance colonoscopy.

As previously indicated, patients who had a prior history of colon cancer or removal of adenomatous polyps could be referred to the SCOPE Program. This pilot program did not change any of the recommendations that were made by colonoscopists at index colonoscopy as it was designed to test several steps in the referral process. In general terms it is often difficult for physicians to change follow-up recommendations made by other physicians, in particular if this would mean that follow-up colonoscopy is postponed to a later date. One of the obvious advantages of having an organized CRC screening program is that follow-up recommendations will be standardized, which will lead to more optimal use of resources.

One of the limitations of our study was the small sample size. Furthermore, the results of our study were collected from a large city in Canada and may not be generalizable to other regions. We also did not analyze the characteristics and practice profiles of the endoscopists in our region, such as number of colonoscopies performed per year. Additionally, since the guidelines have been revised in 2008, differences in practice by clinicians over time may be attributable to adaptation and incorporation of new guidelines and or heightened awareness as CRC screening is becoming more widespread. Our results should be interpreted also knowing there is a lack of an explicit guideline from the Canadian Association of Gastroenterology (CAG). The CAG surveillance guidelines, compared to the AGA guideline, are less explicit and recommendation implies that endoscopists should decide about the appropriate surveillance interval to a greater extent based on clinical judgment. The CAG guideline of 2004 however,

also recommends a 5 year interval for 1 or 2 adenomas and a 3 year interval for 3 or more adenomas similar to the 2008 AGA guideline. Although they do not give specific recommendations for advanced adenomas, they do specify the term and refer to the AGA guideline.

Lastly, the 6-month margin around the optimal follow-up date for colonoscopy was arbitrarily chosen. There are no data in the literature to indicate what an optimal choice is for a time interval around appropriateness. However, we believe the 6-month- interval was a reasonable choice in the context of the current wait times problems for endoscopy in Canada.

In conclusion, a minority of the surveillance colonoscopies are being performed according to the recommendations for surveillance colonoscopy after polypectomy or CRC removal. Where a large proportion of patients who receive surveillance colonoscopies after the detection of adenomas or CRC are seen too often, another group of patients referred for surveillance or screening colonoscopy face a long wait time for gastrointestinal care. The results suggest that efforts should be made to raise awareness among endoscopists about proper surveillance intervals. Our results indicate that quality improvement programs in this area have the potential to result in important clinical benefits for the endoscopy department, especially in the context of wait times and costs.

REFERENCES

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. CA: a cancer journal for clinicians 2011;**61**(2):69-90.
2. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. The American journal of gastroenterology 2008;**103**(6):1541-9.
3. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. The New England journal of medicine 1993;**329**(27):1977-81.
4. Loeve F, van Ballegooijen M, Boer R, et al. Colorectal cancer risk in adenoma patients: a nation-wide study. International journal of cancer 2004;**111**(1):147-51.
5. de Jonge V, Sint Nicolaas J, van Leerdam ME, et al. Systematic literature review and pooled analyses of risk factors for finding adenomas at surveillance colonoscopy. Endoscopy 2011;**43**(7):560-72.
6. Levin B, Lieberman DA, McFarland B, 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. Gastroenterology 2008;**134**(5):1570-95.
7. Leddin D, Bridges RJ, Morgan DG, et al. Survey of access to gastroenterology in Canada: the SAGE wait times program. Can J Gastroenterol 2010;**24**(1):20-5.
8. Leddin D, Armstrong D, Barkun AN, et al. Access to specialist gastroenterology care in Canada: comparison of wait times and consensus targets. Can J Gastroenterol 2008;**22**(2):161-7.
9. Paterson WG, Depew WT, Pare P, et al. Canadian consensus on medically acceptable wait times for digestive health care. Can J Gastroenterol 2006;**20**(6):411-23.
10. Rapuri S, Spencer J, Eckels D. Importance of postpolypectomy surveillance and postpolypectomy compliance to follow-up screening--review of literature. Int J Colorectal Dis 2008;**23**(5):453-9.
11. Mulder SA, Ouwendijk RJ, van Leerdam ME, et al. A nationwide survey evaluating adherence to guidelines for follow-up after polypectomy or treatment for colorectal cancer. J Clin Gastroenterol 2008;**42**(5):487-92.
12. Mysliwiec PA, Brown ML, Klabunde CN, et al. Are physicians doing too much colonoscopy? A national survey of colorectal surveillance after polypectomy. Ann Intern Med 2004;**141**(4):264-71.
13. Krist AH, Jones RM, Woolf SH, et al. Timing of repeat colonoscopy: disparity between guidelines and endoscopists' recommendation. Am J Prev Med 2007;**33**(6):471-8.
14. Davila RE, Rajan E, Baron TH, et al. ASGE guideline: colorectal cancer screening and surveillance. Gastrointest Endosc 2006;**63**(4):546-57.
15. Saini SD, Nayak RS, Kuhn L, et al. Why don't gastroenterologists follow colon polyp surveillance guidelines?: results of a national survey. J Clin Gastroenterol 2009;**43**(6):554-8.
16. Laiyemo AO, Pinsky PF, Marcus PM, et al. Utilization and yield of surveillance colonoscopy in the continued follow-up study of the polyp prevention trial. Clin Gastroenterol Hepatol 2009;**7**(5):562-7; quiz 497.
17. Mulder SA, Van Leerdam ME, Ouwendijk RJ, et al. Attendance at surveillance endoscopy of patients with adenoma or colorectal cancer. Scand J Gastroenterol 2007;**42**(1):66-71.

18. Schoen RE, Pinsky PF, Weissfeld JL, et al. Utilization of surveillance colonoscopy in community practice. Gastroenterology 2010;**138**(1):73-81.

19. Froehlich F, Wietlisbach V, Gonvers JJ, et al. 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.

20. Ben-Horin S, Bar-Meir S, Avidan B. The impact of colon cleanliness assessment on endoscopists' recommendations for follow-up colonoscopy. Am J Gastroenterol 2007;**102**(12):2680-5.

21. Singh H, Nugent Z, Demers AA, et al. Rate and predictors of early/missed colorectal cancers after colonoscopy in Manitoba: a population-based study. Am J Gastroenterol 2010;**105**(12):2588-96.

22. van Kooten H, de Jonge V, Schreuders E, et al. Awareness of postpolypectomy surveillance guidelines: a nationwide survey of colonoscopists in Canada. Can J Gastroenterol 2012;**26**(2):79-84.

23. Sanaka MR, Super DM, Feldman ES, et al. Improving compliance with postpolypectomy surveillance guidelines: an interventional study using a continuous quality improvement initiative. Gastrointest Endosc 2006;**63**(1):97-103.

PART FOUR

Discussion

CHAPTER 9

Summary, discussion, and future perspectives

SUMMARY

Colorectal cancer (CRC) screening is a rapidly moving field. In this thesis we explored strategies and insights that may help to optimize CRC screening programs.

The main topics of the thesis were introduced in Part I. An overview was given of the burden of CRC worldwide and the initiatives to reduce this burden. All modalities eligible for screening were discussed with special focus on fecal tests. These chapters were followed by the aims and outline of the thesis.

In Part II we expanded our knowledge on various strategies in fecal occult blood test-based CRC screening. We compared the diagnostic test accuracy of guaiac fecal occult blood tests and fecal immunochemical tests (FITs) for advanced neoplasia and CRC in average risk individuals and found higher sensitivities for FIT on both outcomes. We demonstrated that detection of advanced neoplasia by FIT increases significantly with age and fecal Hb cut-off concentration. We gave insight in how increasing the cut-off concentration or screening age, reduces colonoscopy demand and affects diagnostic yield of advanced neoplasia and missed lesions. We explored the best screening strategy in terms of number of FIT samples and recommend the use of 1-sample FIT compared to 2-sample FIT screening, because of equal detection of advanced neoplasia after four rounds. We found that, among 1-sample FIT screenees with a fecal hemoglobin (Hb) concentration lower than the applied cut-off, baseline fecal Hb and consecutive fecal Hb concentrations were independent predictors for incident advanced neoplasia risk.

In Part III we showed that quality and readability of online information on CRC screening can be improved. We found that quality standards used in CRC screening programs can also be achieved in routine colonoscopies. Plenary feedback on inter-hospital differences in quality indicators of colonoscopy reduced variation, although overall quality indicators could not be improved. We determined adherence to recommended colonoscopy surveillance following removal of adenomas, and found suboptimal compliance.

This final part discusses the main findings and insights obtained from our research projects and gives directions for future research.

Colorectal cancer (CRC) ranks third among the most commonly diagnosed cancers worldwide, with wide geographical variation in incidence and mortality across the world.[1]

CRC screening tests can be used to identify asymptomatic individuals with advanced adenoma and (early) cancer, together named advanced neoplasia. Removal of these lesions reduces CRC incidence and prevents CRC related mortality.[2-4] There are various screening modalities available for CRC screening, each with advantages and disadvantages.[5] Fecal occult blood tests (FOBT) are widely accepted for this purpose in large-scale population-based screening programs.[6 7] FOBTs can select screenees at higher risk of advanced neoplasia from the large target population and thereby enable the most efficient use of limited colonoscopy resources.

FECAL OCCULT BLOOD TEST BASED COLORECTAL CANCER SCREENING (PART II)

The aim of **Part II** was to explore various aspects of and strategies on FOBT-based screening.

Guaiac fecal occult blood test versus fecal immunochemical test

FOBT screening is based on the principle that a large proportion of colorectal neoplasia bleed microscopically before any clinical signs or symptoms become noticeable for the screenee. Bleeding tends to be intermittent, and blood is distributed unevenly in the stool.[8] The concept of detecting CRC by testing for blood in the stool is based on the observation that cancers bleed because of disruption of the normal mucosa. The amount of blood increases with the size of the polyp and/or the stage of the cancer.[9-12] The guaiac FOBT (gFOBT) was one of the first FOBTs used in CRC screening. Large trials have shown a significant reduction in CRC-related mortality after screening with gFOBT.[2] More recently, the fecal immunochemical test (FIT) has been introduced as an alternative to gFOBT. FIT has several advantages compared to gFOBT screening. These include that FIT detects human blood specifically, in contrast to the gFOBT which can falsely detect other blood substances from digested food. Second, FITs are more specific for lower gastrointestinal tract bleeding since globin is degraded by digestive enzymes in the upper gastrointestinal tract. Third, the sample collection for most FIT variants is less demanding than for gFOBT-sampling, both in terms of number of samples required and handling of stools (smear cards for gFOBTs vs brush/spatula for FIT testing). Fourth, FIT screening does not require dietary

restrictions. Our Cochrane systematic review and meta-analysis showed that FIT had a higher discriminative ability than gFOBT for both advanced neoplasia and CRC as estimated by the HSROC model summary curves (**Chapter 2**). Our findings were strengthened by the fact that most studies were of high methodological quality. Summary sensitivity of FIT for advanced neoplasia and CRC in average-risk individuals was superior to gFOBT. However, the specificity of both tests was similar. Although the majority of clinical trials that demonstrated reduction in CRC mortality used gFOBT, it is successfully argued that using a more accurate FOBT (i.e. FIT) will only improve the mortality benefit from FOBT screening.

These results from **Chapter 2** strongly support current guideline recommendations for implementing FIT-based CRC screening programs or switch from gFOBT to FIT in case of existing programs.[13]

Population-based CRC screening programs that have been implemented over the last years indeed use FIT.[5] For example, The Netherlands implemented a FIT-based national screening program in 2014 with gradual roll-out until 2020.[14] France switched from gFOBT to FIT in 2015.[5] From January 2016, the national screening committee in the UK has recommended a change from gFOBT to FIT after positive results of a FIT pilot study.[15]

Searching for the best FIT screening strategy – age and cut-off

Thus, FIT-screening seems the way to go because of its superior adherence, usability and accuracy. However, within FIT screening different strategies can be applied such as cut-off concentration used and age of the population screened. Both vary between countries, often tailored to available financial resources and colonoscopy capacity.[16 17] A high cut-off and narrow screening age range result in a low positivity rate and consequently low colonoscopy demand.[18] Within the first round of our CRC screening cohort, we estimated the effect of increasing the cut-off concentration and screening age on the colonoscopy demand and advanced neoplasia detection and miss rate (**Chapter 3**).

We found that FIT positivity rates, detection rates and the positive predictive value (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. However, at the expense of missing advanced neoplasia, albeit in a similar ratio for both strategies. The lower sensitivity of higher cutoff concentrations may be compensated by repeating the test in subsequent rounds. Yet, an optimal cut-off concentration or screening age range could not be established in this study in a single screening round. Literature states 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).[19] Since younger persons have more life-years to gain, lower detection rates compared to elder individuals are generally acceptable.

Positivity rate is a surrogate marker for fecal hemoglobin (Hb).[20] Any factor affecting positivity, including age, sex, socio-economic status, previous screening participation or FIT brand will be reflected in the fecal Hb concentration.[21 22] In the Netherlands asymptomatic persons aged 55-75 years are invited biennially for one FIT. At the start of the national screening program, a considerably higher than expected positivity rate was found.[16 23] This was, among other reasons, because the individuals screened were substantially older than in our previous CRC screening trial. To meet the higher than anticipated colonoscopy demand, the cut-off concentration for colonoscopy referral was elevated.[14 16]

Optimizing CRC screening is a continuous process. Screening policymakers should be aware of factors influencing positivity rate and closely monitor outcomes of a screening program to make adjustments if necessary. As the screening program evolves, regular assessment of positivity rates at different fecal Hb cut-offs with distributions by age and sex is recommended.

Searching for the best FIT screening strategy – number of FIT samples

Besides screening age-range and cut-off concentration, different strategies exist with regard to the number of FITs used per screening round. We previously demonstrated that two-sample FIT (2-FIT) screening, using at least one positive test as referral criteria, provides a higher detection rate for advanced neoplasia than 1-sample FIT (1-FIT) screening.[24] However, this was at the expense of higher positivity rates and thus the need for more colonoscopies. Data on participation and diagnostic yield of successive rounds were needed to provide more insight in the long-term effectiveness of 1-FIT versus 2-FIT screening.

We hypothesized that 2-FIT screening might require less screening rounds to be as equally effective as 1-FIT screening in terms of cumulative yield and sensitivity of multiple screening rounds and number of interval cancers. In **Chapter 4** we therefore evaluated the results of 1-FIT versus 2-FIT screening after four screening rounds. Participation is an important early indicator for an effective population-based screening program. In both screening cohorts, a high cumulative participation over four consecutive rounds was seen. Repeated 1-FIT and 2-FIT screening resulted in a similar cumulative yield of advanced neoplasia. The PPV was however lower for 2-FIT than 1-FIT screening, which resulted in a higher (unnecessary) colonoscopy demand.

Interval cancers, i.e. cancers occurring after a negative screening test and before the next screening test is due, are worldwide considered an important indicator of the quality and effectiveness of CRC screening.[25] Although large number of FIT screening studies were performed over the last two decades, few studies reported on interval CRC rates over multiple screening rounds.[26-28] We identified interval CRCs through record linkage with the Dutch Comprehensive Cancer Centre for all participants. The FIT interval CRC rate among screenees participating at least once was 0.11% for 1-FIT and 0.09% for 2-FIT screening, although numbers were too small to reach statistical significance.

Our above described results on 1-FIT and 2-FIT screening at a cut-off concentration of 10 µg Hb/ g feces point towards the use of 1-FIT screening. However, other strategies within 2-FIT screening are possible, such as a higher cut-off concentration, referral criteria of both positive tests (alone or in combination with a higher cut-off), longer screening intervals or combination with other screening tests such as DNA-methylation tests. In our study, screenees with at least one positive FIT at a cut-off concentration of 10 µg Hb/ g feces were referred for colonoscopy and not re-invited. Therefore our study did not allow for simulation of such strategies apart from the first round.[24] Modelling studies are necessary to evaluate (cost-)effectiveness of 1-FIT versus 2-FIT screening by simulating different cut-offs and screening intervals for both strategies. This can further aid in defining the optimal FIT screening strategy.

The ultimate purpose of CRC screening is decreasing mortality, so future studies should report long-term follow up data on mortality and interval cancer rates of FIT-based CRC screening programs. We encourage authors to report a clear definition of interval carcinomas, using standardized nomenclature as developed by the Expert Working Group on interval CRC of the Colorectal Cancer Screening Committee of the World Endoscopy Organization.[25]

Searching for the best FIT screening strategy – use of individual fecal Hb concentrations

FIT has the advantage that fecal Hb concentrations can be measured yielding a quantitative test result. Nonetheless, at present FIT is mostly used like the gFOBT with a pre-determined cut-off for referral to colonoscopy. This strategy does not fully profit from the knowledge that higher fecal Hb concentrations are associated with a higher risk of advanced neoplasia.[18 29-31] Currently, many countries with FIT-based CRC screening programs struggle to match colonoscopy demands with limited resources.[16 17 32] To increase screening efficiency and impact of FIT screening programs, it is relevant to explore if screenees with a negative FIT, (i.e. a fecal Hb concentration below the pre-defined cut-off), can be

categorized according to their actual fecal Hb concentration into different risk groups for development of advanced neoplasia. Such tailored screening would allow for targeted variation of screening intervals, and decrease screening and colonoscopy demand or optimize its use.

Therefore, in **Chapter 5** we aimed to investigate trends in actual fecal Hb concentrations of negative FIT results at first participation and in subsequent rounds as a predictor for future incidence of advanced neoplasia. We showed that fecal Hb concentrations of negative FIT results can be used to predict incident advanced neoplasia and CRC. Furthermore, in this study we were the first to show the added value of fecal Hb concentrations of consecutive negative FIT results over multiple screening rounds in predicting who is at risk of advanced neoplasia. We provided practical heat plots to visualize the risk for screenees after two negative FIT results.

This study supports the use of fecal Hb concentrations in CRC screening optimization. Our findings provide a tool for personalized screening strategies to identify subgroups that are at higher risk of neoplasia within the average-risk target population. This could be similar to coronary heart disease risk management for which cholesterol is a predictor.[33] Changes in cholesterol are monitored over time and combined with other risk factors such as age, gender, body mass index and blood pressure to provide a risk-based score.[33] Future studies should focus on combining fecal Hb concentrations with known CRC risk factors to establish individual tailored screening intervals to maximize screening benefit and efficacy.

QUALITY IN SCREENING AND COLONOSCOPY (PART III)

The aim of **PART III** of this thesis was to provide more insight in factors that are associated with quality in CRC screening, surveillance and colonoscopy.

Optimizing online CRC screening information

The efficacy of CRC screening is dependent on participation and subsequent adherence to surveillance but screening has potential side-effects, including potential harm associated with (follow-up) colonoscopy. The importance of disclosing appropriate information to enable target groups to make informed decisions about screening participation has been emphasized.[34] The internet is increasingly used for health information and therefore the assessment of the availability and quality of online information for screenees is essential.

We evaluated the accuracy, quality, and readability of online information on CRC screening and surveillance in **Chapter 6**. We showed marked variation in accuracy, quality and readability of information on CRC screening websites and that most websites do not address polyp surveillance. Several high quality websites do exist but we learned that it cannot be assumed that screenees will find these websites because of poor correlation between accuracy and Google ranking. Most websites lacked important information regarding polyps and their importance for future surveillance colonoscopies.

Adenomas are found in up to 50% of patients aged 50 years and older who undergo colonoscopy.[35 36] Therefore, also knowledge about post-polypectomy surveillance and post-polypectomy compliance to follow-up screening is important.[37 38] Previous studies have shown that most adenoma patients are unaware of their surveillance recommendations.[39 40] Physicians should explain the need of surveillance, this will likely motivate patients to adhere to surveillance recommendations.

Health care providers interested in developing websites on CRC screening can use our approach to evaluate the quality and readability of provided information to develop the content of the site they are creating. Alternatively they can provide screenees to several of the high quality websites, as identified in **Chapter 6**.

Optimizing colonoscopy

Colonoscopy is the most commonly performed gastrointestinal endoscopic procedure and considered the 'gold standard' investigation of the colon. Aside from primary colonoscopy screening, all CRC screening tests are followed by colonoscopy in case of a positive test result. Measuring and optimizing quality of colonoscopy contributes to an optimal preventive effect of CRC screening and surveillance. The ideal outcome measure for screening is reduction in CRC incidence and mortality. However, this outcome necessitates many subjects and years of follow-up, and therefore surrogate quality measures are used. For this purpose, performance parameters i.e. quality indicators of an individual endoscopist or group of endoscopists are compared with predetermined targets. Despite some limitations, adenoma detection rate (ADR), defined as the proportion of colonoscopies that detect one or more adenomas, is currently considered the best surrogate outcome measure and has emerged by experts as priority quality indicator of colonoscopy.[41] A second priority indicator is adjusted cecum intubation rate (CIR), reflecting the proportion of procedures in which the complete colon is visualized.[41] Low CIR and poor bowel preparation may explain the relative failure of colonoscopy to protect against proximal cancers.[42 43]

Reduction in variation of quality has also emerged as an important priority for colonoscopy practice. We therefore assessed in **Chapter 7** whether plenary feedback on these priority quality indicators can stimulate improvement of routine colonoscopy and thereby reduce inter-hospital differences. We found that quality indicators used in CRC screening programs can also be achieved in routine colonoscopies. In high-standard routine colonoscopy centers, plenary evaluation and discussion of inter-hospital differences reduced variation of quality indicators, although overall quality of colonoscopy was not improved. Higher inter-hospital variation was found for ADR compared to adjusted CIR.

Health care programs should routinely register colonoscopy quality indicators. In this regard, benchmarking variation in quality may highlight additional opportunities to improve care. The use of quality indicators and addressing suboptimal clinical outcomes may reduce variations in colonoscopy performance, and will close the gap between suboptimal and optimal performed colonoscopies. In the Netherlands, colonoscopy quality indicators are incorporated in an accreditation and auditing process, which started in 2014 for endoscopists who participate in the national CRC screening program.[44 45] The Dutch Society for Gastroenterologists is currently working on an registration program for all endoscopists in the Netherlands. This allows endoscopists to measure and benchmark their performance. Endoscopists' knowledge that quality measures are being monitored may in this regard represent a powerful tool that by itself can stimulate improvement of quality of colonoscopy.

Optimizing surveillance

A considerable proportion of all colonoscopies is performed for surveillance purposes. This proportion will further increase with the implementation of CRC screening programs. Adherence to surveillance colonoscopy guidelines is therefore important to prevent unnecessary workload. In **Chapter 8** we evaluated how well Canadian gastroenterologists adhered to colonoscopy surveillance guidelines after adenoma removal or treatment for CRC. Our study showed that in only a third of patients, surveillance colonoscopy was performed at the recommended time interval. Of all surveillance colonoscopies, 46% were performed earlier than recommended. Underuse was also reported, as 21% of the patients received their colonoscopy too late. Deviation from the guidelines did not improve the ADR.

Interventions to improve adherence to surveillance guidelines are needed because inappropriate surveillance colonoscopies seriously hamper the effectiveness and efficiency of surveillance. While a large proportion of patients who receive surveillance colonoscopies after the detection of adenomas or

CRC are seen too often, another group of patients referred for colonoscopy face a long wait time for gastrointestinal care.[46 47] This is concerning because delayed surveillance is associated with an increased rate of advanced adenoma and especially CRC.[48] Also a more recently Italian study reported a high rate of inappropriate recommendations for patients with low risk or no adenomas found at colonoscopy after a positive FIT.[46]

In the light of the recently implemented FIT screening programs for which many countries struggle with colonoscopy capacity, quality improvement initiatives should strive for optimizing the use of surveillance colonoscopy. Efforts should be made to raise awareness among endoscopists about proper surveillance intervals.

REFERENCES

1. GLOBOCAN. Estimated cancer incidence, mortality and prevalence worldwide in 2012 http://globocan.iarc.fr/Default.aspx, 2012.
2. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol 2008;**103**(6):1541-9.
3. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;**366**(8):687-96.
4. Holme O, Bretthauer M, Fretheim A, et al. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. Cochrane Database Syst Rev 2013;**9**:CD009259.
5. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;**64**(10):1637-49.
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. Kuipers EJ, Rosch T, Bretthauer M. Colorectal cancer screening--optimizing current strategies and new directions. Nat Rev Clin Oncol 2013;**10**(3):130-42.
8. Rockey DC. Occult gastrointestinal bleeding. N Engl J Med 1999;**341**(1):38-46.
9. Ciatto S, Martinelli F, Castiglione G, et al. Association of FOBT-assessed faecal Hb content with colonic lesions detected in the Florence screening programme. Br J Cancer 2007;**96**(2):218-21.
10. Edwards JB. Screening for colorectal cancer using faecal blood testing: varying the positive cut-off value. Pathology 2005;**37**:565-68.
11. Levi Z, Rozen P, Hazazi R, et al. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 2007;**146**(4):244-55.
12. Rozen P, Waked A, Vilkin A, et al. Evaluation of a desk top instrument for the automated development and immunochemical quantification of fecal occult blood. Med Sci Monit 2006;**12**(6):MT27-32.
13. Von Karsa LP, J. Segnan, N. (eds). *European guidelines for quality assurance in colorectal cancer screening and diagnosis*. Lyon, 2010.
14. van Veldhuizen HH, M.L. Lansdorp-Vogelaar, I. Adjustment to the implementation of the colorectal cancer screening programme in 2014 and 2015. In: (RIVM) TNIfPHatE, ed. Netherlands: RIVM, 2014.
15. Moss S, Mathews C, Day TJ, 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 2016.
16. van Hees F, Zauber AG, van Veldhuizen H, 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.
17. 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**(3):198-205.
18. 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.

19. Wilschut JA, Hol L, Dekker E, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology 2011;**141**(5):1648-55 e1.

20. Fraser CG, Auge JM, Group P. Faecal haemoglobin concentrations do vary across geography as well as with age and sex: ramifications for colorectal cancer screening. Clin Chem Lab Med 2015;**53**(9):e235-7.

21. Fraser CG, Rubeca T, Rapi S, et al. 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.

22. Kapidzic A, van der Meulen MP, Hol L, et al. Gender Differences in Fecal Immunochemical Test Performance for Early Detection of Colorectal Neoplasia. Clin Gastroenterol Hepatol 2015;**13**(8):1464-71 e4.

23. Penning C, Lansdorp-Vogelaar, I, van Leerdam, ME, van der Meulen, MP, van Vuuren, AJ, Kuipers, EJ, Bonfrer,JM, van Kemenade, FJ, Biermann K, Thomeer, MGJ, Spaander, MCW, Kroep, S, van Ballegooijen, M, de Koning, HJ. Bevolkingsonderzoek Darmkanker: resultaten eerste half jaar 2014, 2014.

24. van Roon AH, Wilschut JA, Hol L, et al. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. Clin Gastroenterol Hepatol 2011;**9**(4):333-9.

25. 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.

26. Steele RJ, McClements P, Watling C, 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.

27. Denters MJ, Deutekom M, Bossuyt PM, et al. 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.

28. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut 2013;**62**(3):409-15.

29. van Doorn SC, Stegeman I, Stroobants AK, et al. Fecal immunochemical testing results and characteristics of colonic lesions. Endoscopy 2015;**47**(11):1011-7.

30. Digby J, Fraser CG, Carey FA, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. J Clin Pathol 2013;**66**(5):415-9.

31. Fraser CG, Mathew CM, McKay K, et al. Automated immunochemical quantitation of haemoglobin in faeces collected on cards for screening for colorectal cancer. Gut 2008;**57**(9):1256-60.

32. Kanavos P, Schurer W. The dynamics of colorectal cancer management in 17 countries. Eur J Health Econ 2010;**10 Suppl 1**:S115-29.

33. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation 1998;**97**(18):1837-47.

34. Wardle J, Robb K, Vernon S, et al. Screening for prevention and early diagnosis of cancer. Am Psychol 2015;**70**(2):119-33.

35. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;**370**(14):1298-306.

36. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;**362**(19):1795-803.

37. Rapuri S, Spencer J, Eckels D. Importance of postpolypectomy surveillance and postpolypectomy compliance to follow-up screening--review of literature. Int J Colorectal Dis 2008;**23**(5):453-9.

38. Schreuders E, Sint Nicolaas J, de Jonge V, et al. The appropriateness of surveillance colonoscopy intervals after polypectomy. Can J Gastroenterol 2013;**27**(1):33-8.

39. Sint Nicolaas J, de Jonge V, Cahen DL, et al. Awareness of surveillance recommendations among patients with colorectal adenomas. Clin Gastroenterol Hepatol 2012;**10**(4):405-11.

40. Schroy PC, 3rd, Lal SK, Wilson S, et al. Deficiencies in knowledge and familial risk communication among colorectal adenoma patients. J Clin Gastroenterol 2005;**39**(4):298-302.

41. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. Am J Gastroenterol 2015;**110**(1):72-90.

42. Singh H, Nugent Z, Demers AA, et al. The reduction in colorectal cancer mortality after colonoscopy varies by site of the cancer. Gastroenterology 2010;**139**(4):1128-37.

43. Rutter MD, Senore C, Bisschops R, et al. The European Society of Gastrointestinal Endoscopy Quality Improvement Initiative: developing performance measures. Endoscopy 2016;**48**(1):81-9.

44. RIVM. Protocol toelating en auditing van coloscopiecentra en endoscopisten, 2012.

45. RIVM en VWS. Landelijke kwaliteitseisen coloscopiecentrum versie 2.1: RIVM; 2013 [Available from: http://www.rivm.nl/dsresource?objectid=rivmp:207707&type=org&disposition=inline&ns_nc=1.

46. Zorzi M, Senore C, Turrin A, et al. Appropriateness of endoscopic surveillance recommendations in organised colorectal cancer screening programmes based on the faecal immunochemical test. Gut 2015.

47. Leddin D, Armstrong D, Borgaonkar M, et al. The 2012 SAGE wait times program: Survey of Access to GastroEnterology in Canada. Can J Gastroenterol 2013;**27**(2):83-9.

48. van Heijningen EM, Lansdorp-Vogelaar I, Steyerberg EW, et al. Adherence to surveillance guidelines after removal of colorectal adenomas: a large, community-based study. Gut 2015;**64**(10):1584-92.

APPENDICES

SUMMARY IN DUTCH

Dikke darmkanker, dat is kanker die ontstaat in de dikke darm of endeldarm, is een belangrijk gezondheidsprobleem. Het staat op de derde plaats van meest voorkomende kankersoorten. Het duurt lang voordat darmkanker zich ontwikkelt en de ziekte geeft pas in een laat stadium klachten. Als darmkanker in een vroeg stadium wordt ontdekt, is de kans op genezing groter en de behandeling minder zwaar. De ziekte begint meestal als een poliep, adenoom genoemd. Adenomen komen regelmatig voor bij mensen ouder dan 50 jaar. Voortgeschreden adenomen hebben een verhoogde kans om uit te groeien tot darmkanker. Als adenomen volledig worden verwijderd kunnen ze niet meer uitgroeien tot darmkanker. Met screening (een bevolkingsonderzoek) kan darmkanker worden voorkomen of in een vroegtijdig stadium worden ontdekt en behandeld. Door invoering van een bevolkingsonderzoek zullen uiteindelijk minder mensen sterven aan deze ziekte. Om deze redenen hebben veel landen een bevolkingsonderzoek naar darmkanker, of wordt deze op dit moment ingevoerd. Er zijn verschillende testen waarmee darmkanker screening uitgevoerd kan worden en voor veel testen zijn meerdere strategieën mogelijk. Dit proefschrift richt zich op strategieën om de kwaliteit van darmkanker screening te optimaliseren.

In het eerste deel (**Deel I**) wordt een uitgebreid overzicht gegeven van wat er tot nu toe bekend is over darmkanker screening, worden de verschillende testen toegelicht en worden de doelstellingen van het proefschrift beschreven. De testen die op dit moment beschikbaar zijn voor darmkanker screening kunnen grofweg worden ingedeeld in ontlasting testen en testen die de binnenkant van de darm afbeelden. De ontlasting test wordt thuis uitgevoerd en vervolgens opgestuurd naar een laboratorium. Daar wordt gekeken of er bloed in de ontlasting zit. Deze hoeveelheid bloed is vaak te weinig om met het blote oog te kunnen zien. Bloed kan wijzen op darmkanker of darmpoliepen die kunnen bloeden als de ontlasting er langs schraapt. Het bloed kan ook komen van bijvoorbeeld aambeien, een scheurtje bij de anus, of een darmontsteking. Om hier uitsluitsel over te geven is een inwendig kijkonderzoek van de darm nodig, een zogenaamde colonoscopie. Hierbij wordt een flexibele slang, waaraan een lampje en een camera zit, via de anus in de dikke darm gebracht.

In het tweede deel (**Deel II**) wordt dieper ingegaan op de ontlastingstest als methode voor darmkanker screening. De twee meest gebruikte ontlastingstesten zijn de guaiac fecaal occult bloed test (gFOBT) en de fecaal immunochemische test (FIT). In **Hoofdstuk 2** hebben we een zogenaamde meta-analyse uitgevoerd, waarbij we alle studies bij elkaar hebben opgeteld die één van deze testen heeft

bestudeerd in de algemene populatie. We vonden dat de FIT darmkanker en voortgeschreden adenomen beter detecteert dan de gFOBT. De FIT lijkt dan ook de meeste geschikte test te zijn om te gebruiken voor darmkanker screening. Onderzoek is nodig hoe het gebruik ervan kan worden geoptimaliseerd. Er zijn namelijk vele strategieën mogelijk om FIT te gebruiken. Zo kan bij sommige FITs de hoeveelheid bloed (fecaal hemoglobine (Hb)) die wordt gemeten in de ontlasting uitgedrukt worden in een getal, een zogenoemde kwantitatieve bepaling. Vervolgens kan bepaald worden boven welke Hb concentratie wordt doorgestuurd voor colonoscopie, de afkapwaarde voor een positieve test. Veel landen worstelen met een tekort aan colonoscopie capaciteit waardoor de screening strategie en afkapwaarde zo gekozen moet worden dat deze aansluit op de beschikbare capaciteit.

In **Hoofdstuk 3** gaven we inzicht in de manier waarop het verhogen van de afkapwaarde of het screenen van een kleinere leeftijdsrange, invloed heeft op de colonoscopie vraag en diagnostische opbrengst van voortgeschreden neoplasie en gemiste laesies. Een hoge afkapwaarde en kleine screening leeftijdsrange leiden tot een laag positiviteitspercentage en dientengevolge een lage colonoscopie vraag. Uit deze analyse bleek verder dat het verhogen van de afkapwaarde of het screenen van een kleinere leeftijdsrange beiden evenredig de colonoscopie vraag en diagnostische opbrengst van voortgeschreden neoplasie en gemiste laesies beïnvloedden.

In **Hoofdstuk 4** verkenden we de beste screening strategie met betrekking tot het aantal FIT testen per screeningsronde. We vergeleken het screenen met 1 FIT per screeningsronde met 2 FITs per ronde. Beide strategieën detecteren na 4 screeningsronden evenveel voortgeschreden neoplasie, terwijl daar voor de 2 FIT strategie meer colonoscopiëen voor nodig zijn. We raden daarom het gebruik van 1 FIT per screeningsronde aan in plaats van 2 FITs per keer.

Na het definiëren van het aantal FITs per ronde, hebben we gekeken naar de mogelijkheden met betrekking tot de fecaal Hb concentratie (hoeveelheid bloed) die wordt gemeten bij het analyseren van de FIT. Op dit moment wordt alleen gekeken of de fecaal Hb concentratie boven of onder de afkapwaarde is waarop doorgestuurd wordt voor colonoscopie. De daadwerkelijke fecaal Hb concentratie wordt buiten beschouwing gelaten. Dit is een gemiste kans omdat in de literatuur al bekend is dat de hoogte van de fecaal Hb concentratie gecorreleerd is met de kans op voortgeschreden adenomen of darmkanker. Daarom hebben we in **Hoofdstuk 5** gekeken of de fecaal Hb concentratie onder de afkapwaarde (dus negatieve FIT) bij de eerste keer deelnemen aan screening voorspellend is op de kans om in de daaropvolgende jaren gediagnosticeerd

te worden met voortgeschreden neoplasie of darmkanker. We toonden aan dat fecale Hb concentraties van negatieve FIT resultaten kunnen worden gebruikt om toekomstige geavanceerd neoplasie en darmkanker te voorspellen. Verder hebben we in deze studie de toegevoegde waarde laten zien van het gebruiken van fecale Hb concentraties van opeenvolgende negatieve FITs om de kans op geavanceerde neoplasie te voorspellen. Dit onderzoek onderschrijft het gebruik van fecale Hb concentraties om darmkanker screening te optimaliseren.

Na uitgebreid te hebben stilgestaan bij de ontlastingstest, belicht **Deel III** van dit proefschrift de kwaliteit van darmkanker screening en colonoscopie.

Colonoscopie is het meest uitgevoerde inwendige onderzoek van het maagdarmkanaal en wordt beschouwd als de 'gouden standaard' voor het onderzoeken van de dikke darm. Het onderzoek wordt niet alleen uitgevoerd na een positieve ontlastingstest, maar bijvoorbeeld ook als personen darmklachten hebben. Indien poliepen worden gevonden en verwijderd worden patiënten vaak opgevolgd (surveillance). In het kader van deze surveillance wordt de colonscopie vaak na enkele jaren herhaald, zodat eventuele nieuw ontstane poliepen ook kunnen worden verwijderd. Het meten en optimaliseren van de kwaliteit van de colonoscopie draagt bij aan een optimale preventieve werking van darmkanker screening en surveillance. In **Hoofdstuk 6** hebben we gedemonstreerd dat de kwaliteit en leesbaarheid van online gegevens over darmkanker screening en surveillance kan worden verbeterd. We evalueerden in **Hoofdstuk 7** de invloed van plenaire terugkoppeling van kwaliteitsparameters van colonoscopie aan zeven Nederlands ziekenhuizen. We vonden dat de kwaliteitsnormen gebruikt voor darmkanker screening colonoscopiëen ook kunnen worden gebruikt voor normale colonoscopiëen. Na plenaire terugkoppeling van de kwaliteitsparameters van colonoscopie nam de variatie tussen ziekenhuizen af. In een groot deel van de ziekenhuizen verbeterde de kwaliteitsparameters van colonoscopie, hoewel de kwaliteitsparameters algeheel niet konden worden verbeterd. In **Hoofdstuk 8** hebben we gekeken naar de naleving van surveillance richtlijnen na verwijdering van adenomen of darmkanker. Het bleek dat de richtlijnen niet goed werden nageleefd, waarbij zowel colonoscopiëen te vroeg werden herhaald, als te laat. Het niet naleven van de richtlijn leidde niet tot een hogere detectie van adenomen.

In het laatste deel **(Deel IV)** worden de belangrijkste bevindingen en inzichten verkregen uit de onderzoeksprojecten samengevat, en geven we op basis daarvan suggesties voor toekomstig onderzoek **(Hoofdstuk 9)**.

ABBREVIATIONS

AA	advanced adenoma
ADR	adenoma detection rate
AN	advanced neoplasia
ASA	American society of anesthesiologists
ASRi	age-standardized incidence rate
ASRm	age-standardized mortality rate
BBPS	Boston bowel preparation score
BCO	Below the cut-off
BMI	body mass index
CCE	colon capsule endoscopy
CI	confidence interval
CIR	cecal intubation rate
CQI	colonoscopy quality indicator
CRC	colorectal cancer
CTC	computed tomography colonography
DTA	diagnostic test accuracy
FIT	fecal immunochemical test for hemoglobin
fHb	fecal hemoglobin
FKG	Flesch-Kincaid Grade Level
FN	false negatives
FP	false positives
FOBT	fecal occult blood test
FRE	Flesch Reading Ease score
FS	flexible sigmoidoscopy
g	gram
gFOBT	guaiac fecal occult blood test
GQS	Global Quality Score
Hb	hemoglobin
HGD	high grade dysplasia
HP	hyperplastic polyp
HR	hazard ratio

IQR	inter quartile range
MAP	mean adenomas per procedure
MAP+	mean adenomas per positive procedure
ml	milliliter
ng	nanogram
OC	optical colonoscopy
OR	odds ratio
PDR	polyp detection rate
PR	positivity rate
PPV	positive predictive value
RCT	randomized controlled trial
SD	standard deviation
TN	true negatives
TP	true postives
TVA	(tubulo) villous adenoma
TA	tubular adenoma
µg	microgram
WAS	website accuracy score

CONTRIBUTING AUTHORS

In alphabetical order
Affiliations at the time this research was conducted.

T.D.G. Belderbos
Department of Gastroenterology and Hepatology
University Medical Center Utrecht
Utrecht, The Netherlands

F. ter Borg
Department of Gastroenterology and Hepatology
Deventer Hospital
Deventer, The Netherlands

P.C.J. ter Borg
Department of Gastroenterology and Hepatology
Ikazia Hospital
Rotterdam, The Netherlands

G.J.M. Borsboom
Department of Public Health
Erasmus MC University Medical Center
Rotterdam, The Netherlands

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

L. van Dam
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

D.M. Le Fèvre
Achmea Health Care
Leiden, The Netherlands

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

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

E.A.M. Hassink
Achmea Health Care
Leiden, The Netherlands

K. Izelaar
Regional organization for Population Screening South-West Netherlands
Rotterdam, The Netherlands

V. de Jonge
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

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

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

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

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

M. Ledeboer
Department of Gastroenterology and Hepatology
Deventer Hospital
Deventer, The Netherlands

M.E. van Leerdam
Department of Gastroenterology and Hepatology
Netherlands Cancer Institute
Amsterdam, The Netherlands

M.A.C. Meijssen
Department of Gastroenterology and Hepatology
Isala Hospital
Zwolle, The Netherlands

L.M.G. Moons
Department of Gastroenterology and Hepatology
University Medical Center Utrecht
Utrecht, The Netherlands

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

R.J.T.H. Ouwendijk
Department of Gastroenterology and Hepatology
Ikazia Hospital
Rotterdam, The Netherlands

L. Rabeneck
Prevention and Cancer Control
Cancer Care Ontario
Toronto, Ontario, Canada

A.H.C. van Roon
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

A. Ruco
Sunnybrook Research Institute
Sunnybrook Health Sciences Centre
Toronto, Ontario, Canada

D. Sadowski
Division of Gastroenterology
University of Alberta Hospital
Edmonton, Alberta, Canada

P. van der Schaar
Department of Gastroenterology and Hepatology
St. Antonius Hospital
Nieuwegein, The Netherlands

R.E. Schoen
Division of Gastroenterology, Hepatology and Nutrition
University of Pittsburgh
Pittsburgh, Pennsylvania, United States of America

P.D. Siersema
Department of Gastroenterology and Hepatology
University Medical Center Utrecht
Utrecht, The Netherlands

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

I. Soo
Division of Gastroenterology
University of Alberta Hospital
Edmonton, Alberta, Canada

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

W.J. Spijker
Regional organization for Population Screening South-West Netherlands
Rotterdam, The Netherlands

E.W. Steyerberg
Department of Public Health
Erasmus MC University Medical Center
Rotterdam, The Netherlands

M. Stouten
Gupta Strategists
Ophemert, The Netherlands

J.J.Y. Sung
Institute of Digestive Disease
Chinese University of Hong Kong
Shatin, Hong Kong, HKSAR, China

T.J. Tang
Department of Gastroenterology and Hepatology
IJsselland Hospital
Capelle aan de IJssel, The Netherlands

S.J.O. Veldhuyzen van Zanten
Division of Gastroenterology
University of Alberta Hospital
Edmonton, Alberta, Canada

W.H. de Vos tot Nederveen Cappel
Department of Gastroenterology and Hepatology
Isala Hospital
Zwolle, The Netherlands

A.J. van Vuuren
Department of Gastroenterology and Hepatology
Erasmus MC University Medical Center
Rotterdam, The Netherlands

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

C. Wong
Division of Gastroenterology
University of Alberta Hospital
Edmonton, Alberta, Canada

G.P. Young
Flinders Centre for Innovation in Cancer
Flinders University
Adelaide, SA, Australia

A.G. Zauber
Department of Epidemiology and Biostatistics
Memorial Sloan Kettering Cancer Center
New York, United States of America

LIST OF PUBLICATIONS

Described in this thesis

1. **E.H. Schreuders**, A. Ruco, L. Rabeneck, R.E. Schoen, J.J.Y. Sung, G.P. Young, E.J. Kuipers. Colorectal cancer screening; a global overview of existing programmes. *Gut. 2015; 64 (10), 1637-1649*

2. **E.H. Schreuders**, E.J. Grobbee, M.C.W. Spaander, E.J. Kuipers. Advances in Fecal Tests for Colorectal Cancer Screening. *Current Treatment Options in Gastroenterology. 2016; 14: 152–162*

3. **E.H. Schreuders***, E.J. Grobbee* , A.H.C. van Roon, L. van Dam, A.G. Zauber, I. Lansdorp-Vogelaar, G.J.M. Borsboom, E.W. Steyerberg, M.E. van Leerdam, M.C.W. Spaander, E.J. Kuipers. Guaiac-based faecal occult blood tests versus faecal immunochemical tests for colorectal cancer screening in average-risk individuals. **both authors contributed equally; Submitted to Cochrane Database of Diagnostic Test Accuracy Reviews*

4. **E.H. Schreuders***, E. Wieten*, S.A.V. Nieuwenburg, B.E. Hansen, I. Lansdorp-Vogelaar, E.J. Kuipers, M.J. Bruno, M.C.W. Spaander. Effects of increasing screening age and fecal hemoglobin cut-off concentration in a colorectal cancer screening program. **both authors contributed equally; Clinical Gastroenterology and Hepatology. 2016 Aug 24 (Epub ahead of print)*

5. **E.H. Schreuders**, E.J. Grobbee, A Kapidzic, A.H.C. van Roon, A.J. van Vuuren, I. Lansdorp-Vogelaar, K. Izelaar M.J. Bruno, E.J. Kuipers, M.C.W. Spaander. Comparison of multiple rounds one versus two-sample fecal immunochemical test-based colorectal cancer screening. *Submitted.*

6. E.J. Grobbee, **E.H. Schreuders**, B.E. Hansen, M.J. Bruno, I Lansdorp-Vogelaar, M.C.W. Spaander, E.J. Kuipers. Faecal haemoglobin concentrations predict future advanced colorectal neoplasia in long-term population-based FIT-screening. *Submitted.*

7. **E.H. Schreuders**, E.J. Grobbee, E.J. Kuipers, M.C.W. Spaander, S.J.O. Veldhuyzen van Zanten. Variable quality and readability of patient-oriented websites on colorectal cancer screening. *Clinical Gastroenterology and Hepatology. 2016 Jul 9. (Epub ahead of print).*

8. **E.H. Schreuders***, T.D.G. Belderbos*, M.A.C. Meijssen, R.J.T.H. Ouwendijk, T.J. Tang, F. ter Borg, P. van der Schaar, D.M. Le Fèvre, M. Stouten, E.A.M. Hassink, W.H. de Vos, P.C.J. ter Borg, M. Ledeboer, M.J. Bruno, L.M.G. Moons, E.J. Kuipers, P.D. Siersema, M.C.W.

Spaander. Plenary feedback on inter-hospital differences to improve variation in quality of colonoscopy **both authors contributed equally; Manuscript in preparation.*

9. **E. Schreuders*,** J. Sint Nicolaas*, V. de Jonge, H. van Kooten, I. Soo, D. Sadowski, c. Wong, M.E. van Leerdam, E.J. Kuipers, S.J. Veldhuyzen van Zanten. The appropriateness of surveillance colonoscopy intervals after polypectomy. ** both authors contributed equally; Can J Gastroenterol. 2013 Jan;27(1):33-8.*

Other studies

10. E. Wieten, **E.H. Schreuders**, E.J. Grobbee, E.J. Kuipers, M.J. Bruno, M.C.W. Spaander. Systematic review and meta-analysis on risk factors of interval CRC in fecal occult blood test based screening. *Manuscript in preparation.*

11. H. van Kooten, V. de Jonge, **E. Schreuders**, J. Sint Nicolaas, M.E. van Leerdam, E.J. Kuipers, S.J. Veldhuyzen van Zanten. Awareness of postpolypectomy surveillance guidelines: a nationwide survey of colonoscopists in Canada. *Can J Gastroenterol. 2012 Feb;26(2):79-84.*

12. D. Kao, E. Lalor, R. Fedorak, G. Sandha B. van der Knoop, S. Doornweerd, H. van Kooten, **E. Schreuders**, W. Midodzi, S. Veldhuyzen van Zanten. A Randomized Controlled Trial of Four Bowel Cleansing Regimens Prior to Colonoscopy. *Can J Gastroenterol. 2011 Dec;25(12):657-62.*

Book chapters

13. E.J. Grobbee, **E.H. Schreuders**, E.J. Kuipers. Are there effective screening programs in Europe? Book chapter in: Valentini V., ed. Multidisciplinary management of rectal cancer, Questions and Answers. 2nd edition. *Springer, 2016 - In press.*

14. E.J. Grobbee, **E.H. Schreuders**, M.C.W. Spaander, E.J. Kuipers. Colorectal cancer diagnosis and screening. Book chapter in: Text book of Hepatogastroenterology. Part II: Gastroenterology. *Jaypee Brothers Medical Pub. 2016 - In press.*

PHD PORTFOLIO

Name PhD student:	E.H. (Eline) Schreuders
PhD period:	January 2013 – December 2016
Erasmus MC department:	Gastroenterology and Hepatology
Promotor:	Prof. dr. Ernst J. Kuipers
Co-promotor:	dr. Manon C.W. Spaander

Courses in methodology and statistics

	Year	Workload
EndNote workshop, Erasmus MC library, Rotterdam	2012	6 hours
Pubmed workshop, Erasmus MC library, Rotterdam	2012	6 hours
Indesign workshop, Molecular medicine postgraduate school, Rotterdam	2013	4 hours
Photoshop & Illustrator workshop, Molecular medicine postgraduate school, Rotterdam	2013	8 hours
Biostatistics for clinicians, Netherlands institute for Health Sciences (NIHES), Rotterdam	2013	40 hours
Cochrane Diagnostic Test Accuracy Review, authors' course Module 3, Cochrane Centre Birmingham, United Kingdom	2013	28 hours
Radiation safety (Stralingsbescherming deskundigheidsniveau 5A en instellingsgebonden Straling hygiënische Regelgeving A), Erasmus MC Zorgacademie	2013	40 hours
Methodology of clinical research and preparation of grant applications, Consultatiecentrum Patiëntgebonden onderzoek (CPO), Erasmus MC, Rotterdam	2013	8 hours
Introduction course on statistics & survival analysis, Molecular medicine postgraduate school, Rotterdam	2013	12 hours
Course Access Basic, Molecular medicine postgraduate school, Rotterdam	2014	6 hours
Writing a Cochrane Diagnostic Test Accuracy Review, Julius UMC Utrecht	2014	16 hours
Introduction in GraphPad Prism, Molecular medicine postgraduate school, Rotterdam	2014	6 hours
WCR2.0 Statistical monthly meetings, Departments of Biostatistics/MDL, Erasmus MC, Rotterdam	2014	12 hours
Biomedical English writing and communication, Erasmus MC, Rotterdam	2014	40 hours

Courses in didactic skills and research integrity

	Year	Workload
Research Management for PhDs/Postdocs, Molecular medicine postgraduate school, Rotterdam	2013	28 hours
BROK-cursus, Consultatiecentrum Patiëntgebonden onderzoek (CPO), Erasmus MC, Rotterdam. Certificate Good Clinical Practice obtained	2013	24 hours
Integrity in scientific research, Dept. of Medical ethics and philosophy, Erasmus MC, Rotterdam	2013	56 hours
Workshop situational leadership, Vrouwen Erasmus MC Netwerk voor Academici (VENA), Erasmus MC Rotterdam	2013	6 hours
Training "Begeleiden van leergroepen", Erasmus MC, Rotterdam	2014	4 hours
Training "Omgaan met groepen voor tutoren", Erasmus MC, Rotterdam	2015	3 hours
Workshop "Coachen van toekomstige Erasmusartsen basis", Erasmus MC, Rotterdam	2015	4 hours
Training "Teach The Teacher – Module I", Erasmus MC, Rotterdam	2015	24 hours
Workshop "Coachen van toekomstige Erasmusartsen vervolg", Erasmus MC, Rotterdam	2016	4 hours

Oral presentations

	Year	Workload
Three rounds of two-FIT screening - World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Vienna, Austria	2014	24 hours
Improvement of colonoscopy quality in daily clinical practice - Nederlandse Vereniging voor Gastro-enterologie najaarscongres, Veldhoven, The Netherlands	2014	12 hours
Repeated two-sample FIT screening for colorectal cancer - Nederlandse Vereniging voor Gastro-enterologie najaarscongres, Veldhoven, The Netherlands	2014	12 hours
Systematic Assessment of Quality of Patient Information on Colorectal Cancer Screening on the Internet - Nederlandse Vereniging voor Gastro-enterologie voorjaarscongres, Veldhoven, The Netherlands	2015	12 hours
Colorectal cancer screening, a global overview of different programs - Symposium Preparing for Screening for CRC in Iceland, Icelandic Cancer Society, Reykjavík, IJsland	2015	24 hours
Screening, wie, hoe en waar? - Renesse Symposium 'Colorectale kankerzorg regisseren?', Renesse, The Netherlands	2016	12 hours
Outcome of one- versus two-sample FIT screening after four rounds - World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Vienna, Austria	2016	24 hours

Poster presentations

	Year	Workload
Third round of two-sample immunochemical fecal occult blood test screening in the Netherlands - United European Gastroenterology week, Vienna, Austria	2014	12 hours
Systematic Assessment of Quality of Patient Information on colorectal cancer screening on the internet - Digestive Disease Week, Washington D.C., United States of America	2015	12 hours
Third round of two-sample immunochemical fecal occult blood test screening in the Netherlands - Digestive Disease Week, Washington D.C., United States of America	2015	12 hours
Plenary feedback on interhospital differences in quality indicators for colonoscopy stimulates quality- Digestive Disease Week, Washington D.C., United States of America	2015	12 hours
Systematic Review of Quality of Patient Information on Colorectal Cancer Screening on the Internet – United European Gastroenterology Week, Barcelona, Spain	2015	12 hours
Four Rounds of Two-Sample Fecal Immunochemical Occult Blood Test Screening. Digestive Disease Week, San Diego, CA, United States of America. Abstract rated top 10% of all abstracts	2016	12 hours
Meta-Analysis on Guaiac-Based Fecal Occult Blood Tests versus Fecal Immunochemical Tests for Colorectal Cancer Screening in Average-Risk Individuals. Digestive Disease Week, San Diego, CA, United States of America.	2016	12 hours
Four Rounds of Two-Sample Fecal Immunochemical Occult Blood Test Screening. United European Gastroenterology week, Vienna, Austria	2016	12 hours

Attended (inter)national conferences

	Year	Workload
Najaarscongres, Nederlandse Vereniging voor Gastro-enterologie, Velhoven, The Netherlands	2013	12 hours
Najaarscongres, Nederlandse Vereniging voor Gastro-enterologie, Velhoven, The Netherlands	2014	12 hours
United European Gastroenterology Week, Vienna, Austria	2014	28 hours
Digestive Disease Week, Washington D.C., United States of America	2015	28 hours
United European Gastroenterology Week, Barcelona, Spain	2015	28 hours
Voorjaarscongres, Nederlandse Vereniging voor Gastro-enterologie, Velhoven, The Netherlands	2015	12 hours
Digestive Disease Week, San Diego, United States of America	2016	28 hours

Attended seminars

	Year	Workload
3e Nationaal congres bevolkingsonderzoek darmkanker, Utrecht, the Netherlands	2013	6 hours
28th Erasmus Liver day. Rotterdam, The Netherlands	2013	6 hours
Wetenschapsmiddag Arts-assistenten vereniging (AAV), Erasmus MC, Rotterdam	2014	6 hours
10th year symposium Gastroenterology, Rotterdam	2015	6 hours
World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Vienna, Austria	2014	8 hours
World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Washington D.C., United States of America	2015	8 hours
World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Barcelona, Spain	2015	8 hours
World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, San Diego, United States of America	2016	8 hours
World Endoscopy Organization (WEO) Colorectal Cancer Screening Meeting, Vienna, Austria	2016	8 hours

Memberships

Netherlands Association of Gastroenterology (NVGE)	2013 – current
Landelijke vereniging van Artsen in Dienstverband (LAD)	2014 – current
American Gastroenterological Assocation (AGA)	2013 – 2015
Vrouwen binnen Erasmus MC Netwerk voor Academici (VENA)	2013 – 2014
Koninklijke Nederlandsche Maatschappij tot bevordering der Geneeskunst (KNMG)	2005 – current

Peer review acitivities

Clinical Gastroenterology and Hepatology

Scottish Government Health Directorates, Translational Clinical Studies Research; grant proposal

Gut

Endoscopy

World Journal of Gastroenterology

Medical Decision Making

Digestive and Liver Disease

International Journal of Cancer

Alimentary Pharmacology & Therapeutics.

Educational activities and lecturing

	Year	Workload
Tutoring first year students curriculum Medicine, Erasmus University Rotterdam, Rotterdam, The Netherlands	2014-2015	40 hours
Supervising graduation project Stella Nieuwenburg, medicine student Erasmus University Rotterdam, Rotterdam, The Netherlands	2014	40 hours
Tutoring first year students curriculum Medicine, Erasmus University Rotterdam, Rotterdam, The Netherlands	2015-2016	40 hours
Gastro-intestinale bloedingen, curriculum spoedeisende hulp en intensvive care verpleegkundigen i.o., Erasmus Zorgacademie	2013-2015	40 hours
Coach professional development ("Professionele ontwikkeling") Bachelor medicine students, Erasmus University Rotterdam, Rotterdam, The Netherlands	2015-2016	24 hours

ACKNOWLEDGEMENTS

Dit proefschrift was niet tot stand gekomen zonder de hulp van velen die (in)direct betrokken zijn geweest, of hebben bijgedragen door simpelweg een onmisbaar onderdeel van mijn leven te zijn. Graag maak ik van de gelegenheid gebruik hen te bedanken.

Allereerst wil ik mijn promotor prof.dr. E.J. Kuipers bedanken. Beste Ernst, het is een grote eer om onder jouw vleugels te mogen promoveren. Je straalt plezier uit die aanstekelijk werkt. Ik bewonder je enthousiasme en doorzettingsvermogen, en hoe je in alle drukte de rust vind om échte aandacht te geven. Ik heb veel van je geleerd, niet alleen over onderzoek, maar ook over andere, net zo belangrijke zaken er om heen, zoals de optimale presentatie van resultaten in manuscripten en presentaties, het communiceren en samenwerken met verscheidene partijen en hoe ver je komt als je altijd eerst ja in plaats van nee zegt.

En dan mijn copromotor dr. M.C.W. Spaander. Lieve Manon, wat was het fijn om zo'n bevlogen copromotor te hebben. Je stimuleerde me haalbare doelen te stellen en meer resultaatgericht te werken, hetgeen voor mij (als perfectionist) soms een uitdaging was. Ik heb genoten van het samen sparren over zaken binnen en buiten het onderzoek. Dankjewel voor je zorgzaamheid en support.

Prof.dr. S.J.O. Veldhuyzen van Zanten, beste Sander, jij liet me thuis voelen in het koude Edmonton en hebt mij enthousiast gemaakt voor wetenschappelijk onderzoek. Ook tijdens mijn promotie-traject bleef je betrokken. Dank voor je doeltreffende wetenschappelijke blik en de gezellige ontbijtjes tijdens congressen. Ik vind het heel bijzonder dat je op deze dag in de commissie wil plaatsnemen.

dr. I. Lansdorp-Vogelaar, beste Iris, veel dank dat je plaats wil nemen in de commissie, en daarnaast voor de fijne samenwerking en je onmisbare analytische blik.

Graag wil ik ook de overige leden van de kleine en grote commissie hartelijk danken voor het beoordelen van mijn proefschrift en het zitting nemen in de commissie te weten; prof.dr. H.J. de Koning, prof.dr. E. Dekker, prof.dr. C Verhoef, prof.dr. C Bangma.

De CRC stuurgroep is de draaiende motor achter het proefbevolkingsonderzoek. Met genoegen denk ik terug aan de waardevolle en doeltreffende stuurgroep vergaderingen waarin, zonder dat hier iets voor nodig was, alle neuzen altijd dezelfde kant opstonden. Dankjewel Kirsten, Wolfert, Hans en Jopie voor jullie ondersteuning vanuit de screening organisatie. Hanneke, Jan, Buddy en Martine, veel dank voor jullie inzet met betrekking tot het analyseren van de ontlastingstesten.

Alle betrokken endoscopisten, endoscopieverpleegkundigen, secretaresses en poli-assistenten van de MDL van het Erasmus MC; bedankt voor de hulp, interesse en prettige samenwerking. Dankjewel Agnes, voor alle logistiek rondom de patiëntenzorg en studies.

De deelnemers aan de wetenschappelijke studies ben ik zeer veel dank verschuldigd. Zonder hen waren veel studies niet mogelijk geweest.

De artikelen in dit proefschrift zouden er niet gekomen zijn zonder de waardevolle inbreng van de coauteurs die ik graag wil bedanken voor de fijne samenwerking en de kennis die we hebben mogen uitwisselen. Dear international co-authors, thank you very much for the fruitful collaboration and your valuable input.

Het is een gek idee dat, alhoewel videocapsule endoscopie een belangrijk onderwerp gedurende mijn promotieproject is geweest, dit boekje hier geen enkel hoofdstuk over bevat. Ik ben erg blij en dankbaar dat dit mooie project alsnog gaat starten. Het ERGO-team met in het bijzonder Bert, Arfan en vooral Jan; we hebben er hard aan getrokken, dankjewel voor de prettige samenwerking en fijne overlegmomenten. Fanny, ik heb het volste vertrouwen in jou als opvolger; heel veel succes!

Lieve mede-promovendi (de dak-duiven en lab-ratten), zonder zo'n fantastische en diverse groep collega's was promoveren niet half zo leuk geweest! Ik denk met een grote glimlach op m'n gezicht terug aan de gezellige koffie-momentjes en borrels, het lief en leed delen met mijn roomies van Ca-419, de bruggenlopen, en de legendarische skireizen en congressen. Het heeft veel mooie vriendschappen opgeleverd.

Lieve paranimfen Els en Esmée, als E-team vormen we een perfect trio. We hebben veel gelachen, gediscussieerd over wetenschap, maar ook gefilosofeerd over zingeving in het leven. Ik bewonder jullie veerkracht en ambitie. Dankjewel dat jullie op deze bijzondere dag naast mij staan.

Een proefschrift schrijven vergt uithoudingsvermogen en maakt persoonlijke support onmisbaar. Dankjewel lieve vriendinnen, door jullie ging ik volgeladen met nieuwe energie naar mijn werk.

Mijn lieve, mooie, zorgzame broertjes Frederic en Maurits en zusje Anne; wat ben ik trots op jullie, precies zoals jullie zijn. En wat geniet ik van het samenzijn met jullie. Lieve papa en mama, ik ben geworden wie ik ben dankzij alle bagage die jullie mij in de loop der jaren hebben meegegeven. Ik zou willen dat ik maar de helft van het vertrouwen in mezelf had dat jullie in mij hebben. Dankjewel voor jullie onvoorwaardelijke liefde.

En dan tot slot; lieve Maarten. Dankjewel dat jij mijn rots in de branding bent. Jij geeft me niet alleen moed om mezelf te zijn, je laat me bloeien. Wat ben ik blij met jou in mijn leven en wat kijk ik uit naar alles wat voor ons ligt.

Eline
Rotterdam, december 2016

ABOUT THE AUTHOR

Elisabeth Helena (Eline) Schreuders was born on August 4, 1986 in Utrecht, The Netherlands. In 2005, she completed her secondary school at 'Christelijk Gymnasium' in Utrecht. In that year she started medical school at the Erasmus University Medical Center in Rotterdam. In 2010, she performed her graduation research at the Division of Gastroenterology, University of Alberta hospital, in Edmonton, Canada. She wrote her graduation thesis under supervision of prof. S.J.O. Veldhuyzen van Zanten on the appropriate use of surveillance colonoscopies in adenoma bearing patients. Her internship in Edmonton resulted in the first published article of this thesis (chapter 8), and it was during this period that she further developed interest in scientific research. In January 2013, she started her PhD trajectory at the department of Gastroenterology and Hepatology of the Erasmus MC University Medical Center under supervision of prof. dr. E.J. Kuipers and dr. M.C.W. Spaander. Her research focused on colorectal cancer screening, which has resulted in this thesis. After finishing her PhD, she aspires to become a general practitioner. She lives in Rotterdam, together with her fiancé Maarten.