IMPLEMENTATION PHASE OF THE DUTCH COLORECTAL CANCER SCREENING PROGRAMME

Esther Toes-Zoutendijk

IMPLEMENTATION PHASE
OF THE DUTCH COLORECTAL CANCER SCREENING PROGRAMME

Esther Toes-Zoutendijk

Colofon

The work presented in this thesis was conducted at the department of Public Health, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

ISBN: 978-94-6361-328-6
Layout and printing: Optima Grafische Communicatie (www.ogc.nl)
Cover design: Femke Molenaar

Financial support for this thesis was kindly provided by the Department of Public Health, Erasmus MC University Medical Center, Rotterdam; Nederlandse Vereniging voor Gastroenterologie; Rijksinstituut voor Volksgezondheid en Milieu.

IMPLEMENTATION PHASE
OF THE DUTCH COLORECTAL CANCER SCREENING PROGRAMME

Implementatiefase
van het Nederlandse bevolkingsonderzoek darmkanker

Proefschrift

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

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 1 november 2019 om 13.30 uur

door

Esther Toes-Zoutendijk

geboren te Dirksland

Erasmus University Rotterdam

Promotiecommissie

Promotor:
Prof. dr. H.J. de Koning

Overige leden:
Prof. dr. G.A. Meijer
Prof. dr. Y.B. de Rijke
Prof. dr. G. van Hal

Copromotoren:
Dr. I. Lansdorp-Vogelaar
Dr. M.E. van Leerdam

CONTENT

1

GENERAL INTRODUCTION

1.1 COLORECTAL CANCER

Disease burden

In 2018 a total of 1,800,977 new cases of colorectal cancer (CRC) were diagnosed worldwide; 1,006,019 in men and 794,958 in women. This makes it the third and second most common cancer in men and women, respectively.[1] Although an increasing incidence is observed in non-Western countries, the largest CRC burden is still present in developed countries. In Europe, 500,000 new cases were reported in 2018, although incidence varied between countries.[2] In the Netherlands, before the introduction of screening, yearly 13,000 individuals were newly diagnosed with CRC.[3] Worldwide, the incidence of CRC is increasing due to ageing of the population, change in dietary habits and rise in risk factors like smoking, obesity, and lack of physical activity.[4,5] It is expected that without interference the number of CRC cases in the Netherlands will increase from 13,000 to 17,000 persons per year by 2020.[6] Life-time risk of developing CRC is 5% for men and 4% for women in the Netherlands.[7] This is slightly lower than the observed life-time risk in the United Kingdom, estimating a life-time risk of 7% for men and 6% for women.[8]

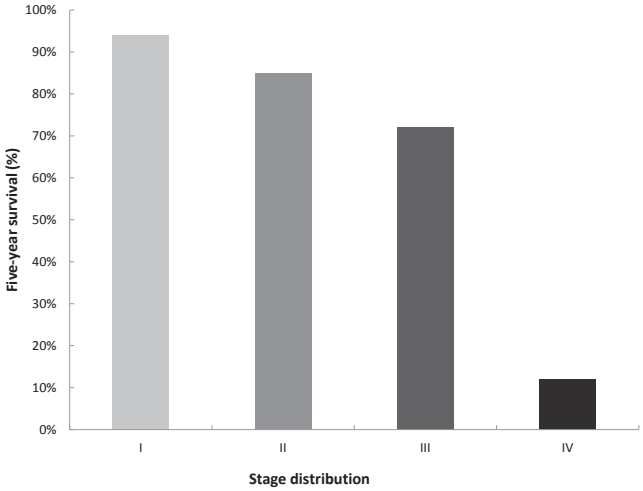
In 2018, worldwide 861,663 persons died of CRC; 474,606 men and 387,057 women. Therewith, it is the fourth and third cancer-related cause of death worldwide in men and women, respectively.[1] In Europe, 243,000 individuals died of CRC in 2018.[2] There is a wide variation in CRC-related mortality rates, with higher mortality rates in less developed countries. Consequently, 5-year survival varies from 35% in Poland to 58% in Finland and 60% in Sweden.[9,10] This variation is probably the result of different cancer treatment or stage distribution at diagnosis. Recent numbers from the Netherlands showed a 5-year relative survival of 61%.[7] The high incidence and mortality rates indicate that CRC is a major health problem.

Survival strongly depends on cancer stage at time of diagnosis.[11,12] Staging of CRCs is done according to the 7th edition of the TNM classification.[13] Stage 0 is considered carcinoma in situ. Stage I are tumours that were confined to the submucosa or had grown into the muscularis propria. Stage II are tumours that have invaded the serosa or penetrated to the peritoneal surface or other organs but without locoregional lymph node involvement. Stage III are tumours that also have metastasis in the locoregional lymph nodes. Stage IV are tumours that have distant metastases. In the Netherlands, 5-year survival is 94% for stage I compared to 12% for stage IV (Figure 1).[7] This strong association between survival and stage distribution emphasises the importance to detect CRCs as early as possible, as this will improve survival after CRC diagnosis.

Progression of colorectal cancer

Development from a small polyp into CRC is characterised by a multistep process involving series of histological, morphological, and genetic changes over time. Currently, two CRC

Figure 1: Five-year survival by stage distribution of individuals with colorectal cancers in the Netherlands



pathways have been identified. The first, so-called traditional pathway, gives rise to 70-85% of all CRC.[14] In this pathway, the normal colon epithelial cells change into aberrant crypt foci, and subsequently into small non-advanced adenomas (<1cm in size, with tubular histology). These adenomas can progress into advanced adenomas (AA) (adenomas with histology showing >=25% villous component or high-grade dysplasia or size >= 10 mm). From AA it can develop into early cancers and lastly advanced cancers with an accumulation of somatic mutations.[15,16] Besides this conventional adenoma pathway, there is an alternative pathway, the so-called serrated neoplasia pathway. This pathway has another precursor lesion, the serrated polyp. It is estimated that 15-30% of the CRCs results from this pathway.[17] Serrated lesions are divided in three subgroups: hyperplastic polyps, sessile serrated polyps and traditional serrated polyps. Of these subtypes, hyperplastic polyps are thought not to develop into CRCs.

As described above, CRC disease is characterised by a long pre-malignant stage. The dwell time is the time from the development of adenomas to symptom-detected CRCs in the absence of screening, which is estimated with microsimulations models to be 17-25 years.[18] The pre-cancerous stage polyps, either early adenomas or sessile serrated lesions, are asymptomatic. With advancing lesions, symptoms may become present but are often a-specific: abdominal pain, change in bowel habits, rectal blood loss, or weight loss.[19] By the time the signs of CRC become evident, the disease has often already developed in an advanced stage with poor associated survival rates.

Aetiology

Table 1 shows an overview of several risk factors and their impact on the development of CRC.[20] These factors can be divided in two subgroups, modifiable and non-modifiable risk factors:

Table 1: Overview of risk and preventive factors of colorectal cancer
Adapted from Brenner et al.[20] with permission.

	Risk
Sociodemographic factors	
Older age	↑↑↑
Male sex	↑↑
Medical factors	
Family history	↑↑
Inflammatory bowel disease	↑↑
Diabetes	↑
Helicobacter pylori infection	(↑)
Other infections	(↑)
Colonoscopy	↓↓
Hormone replacement therapy	↓
Aspirin	↓
Statins	(↓)
Lifestyle factors	
Smoking	↑
Excessive alcohol consumption	↑
Obesity	↑
Physical activity	↓
Diet factors	
High consumption of red and processed meat	↑
Fruit and vegetables	(↓)
Cereal fibre and whole grain	(↓)
Fish	(↓)
Dairy products	(↓)

↑↑↑=very strong risk increase. ↑↑=strong risk increase.↑=moderate risk increase.
↓↓=strong risk reduction. ↓=moderate risk reduction.
Parentheses show probable but not fully established associations.

Modifiable risk factors

Different modifiable factors can lead to an increased risk for CRC. First, choice of diet can impact your risk for CRC. It has been shown that intake of processed meat or red meat increases the risk for CRC up to 17-18%.[21] Note, large amounts of red meat have to be consumed (100 g/day). Second, obesity, low levels of physical activity, alcohol consumption

and cigarette smoking are also be related with an increased risk for CRC.[22,23] In contrast, intake of calcium, whole grains, fibre, and fruit and vegetables might decrease the risk for CRC up to 50%.[24,25] It was estimated that 45% of all CRCs were attributable to an unhealthy lifestyle, irrespectively of a person's genetic risk.[26] Therefore, a healthy lifestyle with physical activity and healthy diet might lower the risk of CRC.

Non-modifiable risk factors

There are various non-modifiable factors that increase individuals CRC risk. Well-known non-modifiable risk factors are sex and age.[1,27] Besides these two important risk factors, several diseases can lead to an increased CRC risk. Some examples are inflammatory bowel disease, type II diabetes and cystic fibrosis.[28-31] Lastly, DNA plays an important role in the development of CRC. Genetic contribution to CRCs can be divided in a few subgroups: family history with nonhereditary CRC, hereditary CRC syndrome (such as Lynch syndrome and familial adenomatous polyposis (FAP)), and other genetic variation (known as single-nucleotide polymorphisms (SNPs)).[27,32,33]

Modifiable and non-modifiable risk factors should be considered together to determine the overall risk for CRC. The combination of family history, environmental factors and genetics on top of age and gender will give the best prediction for an individual's CRC risk.[27] It is unknown whether the impact of the above-mentioned risk factors is similar for the two precursors of CRC: conventional adenomas and serrated polyps. A recent study suggests that both precursors share most common risk factors, but the magnitude of the association might differ.[34] Cigarette smoking, BMI, and alcohol consumption were more strongly associated with serrated polyps, whereas physical activity and dietary factors like folate, calcium, and Vitamin D had a stronger inverse association with conventional adenomas.

1.2 COLORECTAL CANCER PREVENTION

Reducing the burden of CRC could be established in three ways: primary prevention, secondary prevention and tertiary prevention. As the focus of the thesis is on screening, secondary prevention will be explained in more detail.

Primary prevention

It was estimated that almost half of all CRCs are attributable to an unhealthy lifestyle, such as smoking, alcohol consumption, diet, limited physical activity and body fatness.[26] Therefore, it is of great importance that primary prevention will be focussed on these risk factors. Additional benefit of reducing these risk factors is the positive side effects on many other diseases such as diabetes and cardiovascular diseases.

Besides a healthy lifestyle there is some evidence for a preventive effect of certain drugs (chemoprevention) on CRC. Best-known chemoprevention agents are aspirin and nonsteroidal anti-inflammatory drugs (NSAIDs).[35] Although aspirin and NSAIDs have the potential to lower the CRC risk, they may also cause negative side effects, such as haemorrhagic strokes and gastrointestinal complications such as peptic ulcers and bleeding. Maybe it should only be considered for specific high-risk groups.[20,36] However, there is no guideline yet on the usage of chemoprevention agent.

Secondary prevention

With screening, asymptomatic individuals are systematically tested to identify the disease or risk factors for the disease. Screening can prevent the disease or detect the disease in an earlier stage. CRC is a good candidate for screening as it is a slow growing cancer, characterised by a long pre-malignant disease stage. Premalignant lesions can be removed before they become cancer or otherwise CRCs can be detected in an early stage.[11,12] Adenomas with histology showing >=25% villous component or high-grade dysplasia (both considered as AA), are in generally larger in size and are more likely to conceal cancer cells. Also, the risk for adenomas to develop into CRC increases as the size of the polyp increases (>10mm, also considered as AA).[37] Accordingly, AA is also considered as relevant finding of CRC screening. Both CRC and AA are therefore considered as true positives in CRC screening.

There are many different screening methods available for CRC screening. The most commonly used screening methods in Europe are stool-based occult blood test and endoscopy methods. Two types of stool-based occult blood test are used, the guaiac Faecal Occult Blood (gFOBT) and the faecal immunochemical test (FIT). Most important difference between those two tests is that FIT is a quantitative test, enabling to choose the preferred cut-off (µg Hb/g faeces) for referral for follow-op colonoscopy. This is important when considering a desired balance between true and false positive test results or encountering colonoscopy capacity problems. Two endoscopy methods are carried out, sigmoidoscopy and colonoscopy.[38,39] Stool-based test and sigmoidoscopy also have to be followed by a colonoscopy for diagnosis and removal of lesions. Computed tomography colonography (CTC), so-called virtual colonoscopy, is another CRC screening method. This screening method is hardly offered within an organised programme. Newer screening methods are also available, like multitarget-stool DNA testing, SEPT9 biomarker assay or video capsule endoscopy.[40-42] These newer screening methods are currently not offered in population-based screening programmes in Europe.

There is robust evidence that both repeated gFOBT and once-only flexible sigmoidoscopy screening can reduce CRC-related mortality.[43-49] No evidence is available from randomised controlled trials of the FIT on mortality reduction. gFOBT and FIT are similar tests, both stool-based tests; however performance of FIT is superior to gFOBT. Therefore, it is expected that the mortality reduction with FIT could even be larger than gFOBT. Additionally, there is

evidence on mortality reduction from observational studies.[50-54] Therefore, it is assumed that FIT screening will also result in a reduction of CRC-related mortality. Currently randomised controlled trials of colonoscopy screening are executed. Estimates of long-term effect of colonoscopy screening on CRC-related mortality will soon be available.[55-56] Colonoscopy is similar to sigmoidoscopy, but inspects the entire colon whereas sigmoidoscopy only inspects the lower part of the colon. As colonoscopy screening has a better test performance than sigmoidoscopy screening, combined with the evidence from observational studies, reduction of CRC-related mortality is also expected.[57] Note, to observe a mortality reduction within the population, it is crucial that not only a proper screening test is used, but also that the screenings test is accepted within the population.

Another important aspect that should be considered before implementing CRC screening is the harm-benefit ratio. Benefits of screening have been discussed above, the potential of various screening methods to reduce CRC-related mortality. However, there are more noteworthy benefits like reduction in advanced disease stage and reduction of the CRC incidence. The harms of screening differ substantially between screening methods. Harms of stool-based test could be psychological distress after receiving a positive test result and fear of CRC diagnosis.[58,59] The aversion of individuals to perform a stool test may also be considered as harm.[60] An important harm is the number of false-positives undergoing an unnecessary follow-up with colonoscopy. Fear of receiving a positive test result and CRC diagnosis also apply for endoscopy screening. However, endoscopy screening can have more substantial harms than FOBT screening, namely endoscopy-related complications. Estimated risk for a major bleeding was estimated to be 8 per 10,000 colonoscopies and for a perforation 4 to 7 per 10,000 colonoscopies.[59,61] Fatal complications after colonoscopy are very rare. Meta-analyses estimated the mortality rate ranging from 3 to 7 deaths per 100,000 colonoscopies.[62] Note, this rate includes fatal complications of colonoscopy with all indications and therefore not directly applicable on endoscopy screening. Another harm associated with screening, regardless of the screening method, is overdiagnosis. Overdiagnosis in CRC screening concerns detection of polyps or CRCs that, without screening, would not have been diagnosed in an individual's lifetime. It is unknown which polyp's progress or deteriorates. Therefore, it is uncertain which polyps would never develop into CRC and therefore will be removed unnecessarily. Thus, it is uncertain to what extent overdiagnosis is present in CRC screening. But inviting older individuals or individuals with comorbidities to participate in FIT screening, will most likely lead to overdiagnosis.[63] However, quantification of the magnitude of overdiagnosis in CRC screening is currently lacking. It is very complicated to come up with a good estimation, as by the removal of precancerous polyps CRCs will also be prevented. To sum up, CRC screening is associated with harms, however serious harms like death as a result of endoscopy rarely occur. In generally, it is considered that the benefits of CRC screening outweigh the harms.[52]

Tertiary prevention

Treatment of individuals with a CRC diagnosis to prevent further complications is considered as tertiary prevention, so-called survivorship. Tertiary prevention aims to prevent further health impact and improve quality of life after a diagnosis with CRC. Treatment of CRC is continuously changing with new innovations. A recent Dutch study presented an overview of the last 25 years of CRC treatment.[64] This study showed an increase in the use of postoperative chemotherapy for individuals diagnosed with stage III colon cancer and an increase in preoperative radiotherapy for rectal cancer. Another increase was the more intensified care of stage IV CRC, resulting in improved outcomes.[64] Other preventive strategies besides CRC treatment are similar to primary CRC prevention. But the target is on treatment-related side-effects or CRC-related morbidity. Strategies for tertiary prevention besides treatment options are an understudied topic. Known examples of strategies are; physical activity, healthy diet containing vitamin D, fibre, coffee, marine omega-3 fatty acid. Those strategies might improve survival and quality of life.[25,65]

1.3 MONITORING AND EVALUATION OF SCREENING PROGRAMMES

Screening programmes

Worldwide, many countries have implemented a CRC screening programme.[38] Programmes are predominantly introduced in high income countries. Screening can be designed as opportunistic or organised programmes. In an organised screening programme, like in the Netherlands, the entire target population receives an invitation to participate. In an opportunistic screening programme, like in the US and Germany, screening is recommended and reimbursed but depends on individuals' decision. They have to request the screening test themselves at the doctor or pharmacy.

Choosing the best screening strategy for the population is a complex process. When deciding on which test to use several aspects should be taken into account: test sensitivity, specificity, population preference, adherence, harms, capacity and costs. Colonoscopy has the highest sensitivity of all CRC screening methods; however it has downsides like severe complications, high costs, lack of adherence and straining colonoscopy capacity.[66,67] All these downsides need also to be considered when offering screening to the total population. There is a growing recognition that an optimal screening method heavily depends on population preference and availability of resources.[68,69] Besides the choice for the best test, starting age, stopping age and screening interval should be explored to design most effective screening programme for a population.

The Dutch colorectal cancer screening programme

The Netherlands may serve as an excellent example weighing all these various aspect of screening in the decision for the optimal screening method for the Dutch population. In the Netherlands an extensive preparatory process has taken place before the implementation of the national population-based CRC screening programme.[70] This process started with a report from the Dutch Health council in 2001 indicating the need for a national screening programme.

Pilot studies

In 2006 pilot studies were initiated to study the potential of a national CRC screening programme in the Netherlands.[51,71-74] The aim of these Dutch pilot studies was to evaluate most important aspects (i.e. participation, diagnostic yield and cost-effectiveness) of most relevant screening methods: gFOBT, FIT (with FIT cut-offs ranging from 10-40 µg Hb/g faeces), colonoscopy, sigmoidoscopy and CTC. These trials were conducted in Rotterdam, Amsterdam and Nijmegen. FIT screening showed the highest participation rate up to 60-62% in the first round, compared to 47-50% for gFOBT, 32% for sigmoidoscopy, 34% for CTC and 22% for colonoscopy. FIT also showed the highest diagnostic yield, with the highest detection of CRC per 1,000 invitees over two screening rounds.[75] Because of these favourable outcomes of FIT screening, the next step was to determine the optimal FIT cut-off. Outcomes of the pilot studies and subsequent modelling were used to inform policy makers to decide on the most optimal or feasible cut-off for referral to colonoscopy follow-up.[70]

Modelling studies

Microsimulation Screening Analysis (MISCAN)-Colon, a decision model that can be used to predict the benefits, harms and associated costs of different CRC screening strategies, was used to determine the optimal FIT cut-off. This model showed that a FIT cut-off of 10 µg Hb/g faeces will result in highest sensitivity and will be most effective.[76] With unlimited colonoscopy the optimal screening strategy for the Dutch population would be an annual FIT, with a cut-off of 10 µg Hb/g faeces for individuals aged 45-80 years.[77] But in practice, colonoscopy capacity is not unlimited. The model demonstrated that with restricted colonoscopy capacity, the most effective strategy would be annual screening with a FIT cut-off of 40 µg Hb/g faeces and smaller age range of individuals aged 50-75 years.[77]

Health Council

The Health Council plays an important role in designing and implementing a national screening programme in the Netherlands, advising the Minister of Health. They strongly advised on biennial screening with a FIT cut-off of 15 µg Hb/g faeces for individuals aged 55-75 years old.[78] This advice was based on the outcomes of the pilot studies and subsequent

modelling, but also on expert opinion and literature review. The final screening strategy in terms of FIT cut-off, interval and age range was based on the several considerations.

A FIT cut-off of 15 µg Hb/g faeces was advised by the Health council because it has a more favourable balance between true-positives and false-positives (higher positive predictive value (PPV)) and it results in a lower colonoscopy demand. Increasing the FIT cut-off from 10 to 15 µg Hb/g faeces will have a minor impact on CRC detection, but more AA and non-advanced adenoma will be missed.[71]

Annual gFOBT did show a higher CRC-related mortality reduction compared to biennial gFOBT screening (33% versus 20%).[47] However, additional benefit of annual screening over biennial screening is debated.[79] The Health Council concluded that the disadvantage of screening every year, as opposed to every second year, is that screening cost will almost be twice as high, while the desirable effects increase by smaller amounts. Therefore, biennial screening was considered as a more attractive option. The Health council therefore advised that the extra costs involved in annual screening do not outweigh the potential extra benefits.

The target age group was narrowed to individuals aged 55-75 years. This higher starting age was chosen because of the lower incidence of CRC in younger individuals at that time.[78] The lower stopping age was decided to avoid the higher risk of colonoscopy-related complications in older individuals. This recommendation was based on the results of the modelling studies of our research group.

Design Dutch organised CRC screening programme

In accordance with the advice of the Health council, the minister of Health decided on May 25, 2011 to gradually implement a national population-based screening programme with biennial FIT with a cut-off of 15 µg Hb/g faeces for men and women aged 55 to 75 years. The Dutch CRC screening programme was gradually implemented by age groups from 2014 onwards. This phased implementation of five years allowed a timely increase of the colonoscopy capacity. Ultimately, in 2019 all individuals between 55-75 years old should have been invited at least once.

Relevance of monitoring and evaluation

The European guidelines for quality assurance in CRC screening state the relevance of monitoring and evaluation as follows: evaluation and interpretation of screening outcomes are essential to recognise whether a CRC screening programme is achieving the goals for which it has been established.[80] Twenty important recommendations on CRC screening are given in this extensive guideline. Examples of relevant recommendations are: database with individual's records, annual monitoring reports by age and gender, minimal FIT participation of 45%, minimal participation to follow-up colonoscopy of 90%, more favourable stage distribution for screen-detected CRCs than symptom-detected CRCs, and evaluation of

interval CRCs. If the above-mentioned recommendations are followed, it is expected that CRC screening will be effective in reducing CRC-related mortality.

In the Netherlands, as described above, an extensive preparatory process was followed before the implementation of a national CRC screening programme. Next, during a planning phase public tenders for test, laboratories and packaging were set out. In addition, quality assurance and accreditation programmes were set up for endoscopy, pathology and laboratories. Also, a large IT infrastructure was developed. This information system automatically structures the total screening process, integrates information from different sources and is continuously updated. This national information system (ScreenIT) enables real-time monitoring of the national CRC screening programme. Monitoring of a screening programme is crucial. Although the design of the Dutch CRC screening programme is evidence-based and well-planned, it is unknown if the performance on a national level will be in line with expectations. The Netherlands is indeed a good example that expectations and reality are not in line. In the first year in 2014 weekly monitoring was carried out to evaluate the performance of the national CRC screening programme. These weekly reports showed high positivity rates, low PPV and an increase in waiting period for colonoscopy. Consequently, the programme was adjusted after 6 months that will be described in more detail in Chapter 2.

1.4 AIM AND RESEARCH QUESTION

The general aim of this thesis is to evaluate the implementation phase of the national CRC screening programme in the Netherlands. After many years of extensive preparations, expectations on programme performance were high. To ensure that these expectations are met on a national level, monitoring and evaluation of the national screening programme in real setting c.q. national level is important. This led to the following research question of this thesis:

Is the performance of the Dutch colorectal cancer screening programme during the implementation phase satisfying and according to expectations?

The performance of the Dutch CRC screening programme was evaluated separately per performance indicator; participation FIT, FIT positivity, participation to follow-up colonoscopy, CRC and AA detection, stage distribution, location, interval cancers, socioeconomic differences and consistency of FIT performance. The outcomes of important performance indicators were also compared with surrounding countries using FIT screening.

1.5 OUTLINE OF THIS THESIS

Chapter 2 to 7 addresses the Dutch national population-based CRC screening programme. In **Chapter 2** the first year of the Dutch national population-based CRC screening programme was evaluated. It describes the relevance of real-time monitoring to optimise programme performance. We evaluated the participation rate, positivity rate, PPV and detection rates before and after needed adjustment of the FIT cut-off. **Chapter 3** we compared the stage distribution of screen-detected and symptom-detected CRCs. In **Chapter 4** the programme performance of the second screening round was evaluated. We also estimated the impact of the adjusted FIT cut-off on positivity rate, PPV and detection rates. In In **Chapter 5** we estimated the interval CRC incidence and FIT sensitivity after the first screening round and the impact of the adjusted FIT cut-off. In **Chapter 6** we evaluated social economic status (SES) differences in participation and yield of FIT screening. We used area SES and compared the performance indicators participation rate, positivity rate, PPV and detection rate. In **Chapter 7** we evaluated the consistency of FIT in testing positive or detecting CRC or AA for different batches of specimen collection devices, lot reagents and laboratories. **Chapter 8** compared important screening programme indicators of four organised CRC screening programmes using FIT; Basque country (Spain), France, Flanders (Belgium) and the Netherlands. In the general discussion in **Chapter 9**, the research question will be answered and discussed per element. Subsequently the methodological considerations of the analyses in this thesis will be explained and future perspectives will be touched on briefly. Lastly, overall conclusions will be drawn and recommendations will be given.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019;144(8):1941-53.
3. Elferink MA, van der Vlugt M, Meijer GA, et al. Colorectal carcinoma in the Netherlands: the situation before and after population surveillance. Colorectaal carcinoom in Nederland: situatie voor en na invoering van het landelijke bevolkingsonderzoek. Ned Tijdschr Geneeskd. 2014;158:A7699.
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. Pilleron S, Sarfati D, Janssen-Heijnen M, et al. Global cancer incidence in older adults, 2012 and 2035: A population-based study. Int J Cancer. 2019;144(1):49-58.
6. KWF Kankerbestrijding. Rapport kanker in Nederland tot 2020: Trends en prognoses. 2011. https://www.kwf.nl/sitecollectiondocuments/rapport-kanker-in-nederland-tot-2020.pdf.[Cited February 5 2019].
7. Nederlandse Kanker Registratie (NKR). http://cijfersoverkanker.nl. [Cited February 5 2019].
8. Bowel cancer incidence statistics. www.cancerresearchuk.org. [Cited February 5 2019].
9. Gatta G, Trama A, Capocaccia R. Variations in cancer survival and patterns of care across Europe: roles of wealth and health-care organization. J Natl Cancer Inst Monogr. 2013;2013(46):79-87.
10. Morris EJ, Sandin F, Lambert PC, et al. A population-based comparison of the survival of patients with colorectal cancer in England, Norway and Sweden between 1996 and 2004. Gut. 2011;60(8):1087-93.
11. O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst. 2004;96(19):1420-5.
12. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65(2):87-108.
13. Edge SB, Byrd DR, Compton CC, et al. editors. AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer- Verlag, 2009.
14. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology. 2010;138(6):2088-100.
15. Jones S, Chen WD, Parmigiani G, et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci U S A. 2008;105(11):4283-8.
16. 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.
17. Dekker E, IJspeert JEG. Serrated pathway: a paradigm shift in CRC prevention. Gut. 2018;67(10):1751-2.
18. van Hees F, Habbema JD, Meester RG, et al. Should colorectal cancer screening be considered in elderly persons without previous screening? A cost-effectiveness analysis. Ann Intern Med. 2014;160(11):750-9.
19. Cappell MS. Pathophysiology, clinical presentation, and management of colon cancer. Gastroenterol Clin North Am. 2008;37(1):1-24, v.
20. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014;383(9927):1490-502.
21. Bouvard V, Loomis D, Guyton KZ, et al. Carcinogenicity of consumption of red

and processed meat. Lancet Oncol. 2015;16(16):1599-600.
22. Haggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg. 2009;22(4):191-7.
23. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. Nat Rev Dis Primers. 2015;1:15065.
24. Santarelli RL, Pierre F, Corpet DE. Processed meat and colorectal cancer: a review of epidemiologic and experimental evidence. Nutr Cancer. 2008;60(2):131-44.
25. Song M, Garrett WS, Chan AT. Nutrients, foods, and colorectal cancer prevention. Gastroenterology. 2015;148(6):1244-60 e16.
26. Carr PR, Weigl K, Jansen L, et al. Healthy Lifestyle Factors Associated With Lower Risk of Colorectal Cancer Irrespective of Genetic Risk. Gastroenterology. 2018;155(6):1805-15 e5.
27. Jeon J, Du M, Schoen RE, et al. Determining Risk of Colorectal Cancer and Starting Age of Screening Based on Lifestyle, Environmental, and Genetic Factors. Gastroenterology. 2018;154(8):2152-64 e19.
28. Jess T, Rungoe C, Peyrin-Biroulet L. Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. Clin Gastroenterol Hepatol. 2012;10(6):639-45.
29. Jess T, Simonsen J, Jorgensen KT, et al. Decreasing risk of colorectal cancer in patients with inflammatory bowel disease over 30 years. Gastroenterology. 2012;143(2):375-81 e1; quiz e13-4.
30. Overbeek JA, Kuiper JG, van der Heijden A, et al. Sex- and site-specific differences in colorectal cancer risk among people with type 2 diabetes. Int J Colorectal Dis. 2019;34(2):269-76.
31. Billings JL, Dunitz JM, McAllister S, et al. Early colon screening of adult patients with cystic fibrosis reveals high incidence of adenomatous colon polyps. J Clin Gastroenterol. 2014;48(9):e85-8.
32. Henrikson NB, Webber EM, Goddard KA, et al. Family history and the natural history of colorectal cancer: systematic review. Genet Med. 2015;17(9):702-12.
33. Kastrinos F, Stoffel EM. History, genetics, and strategies for cancer prevention in Lynch syndrome. Clin Gastroenterol Hepatol. 2014;12(5):715-27; quiz e41-3.
34. He X, Wu K, Ogino S, et al. Association Between Risk Factors for Colorectal Cancer and Risk of Serrated Polyps and Conventional Adenomas. Gastroenterology. 2018;155(2):355-73 e18.
35. Din FV, Theodoratou E, Farrington SM, et al. Effect of aspirin and NSAIDs on risk and survival from colorectal cancer. Gut. 2010;59(12):1670-9.
36. Rothwell PM, Fowkes FG, Belch JF, et al. Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. Lancet. 2011;377(9759):31-41.
37. Simon K. Colorectal cancer development and advances in screening. Clin Interv Aging. 2016;11:967-76.
38. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49.
39. Senore C, Basu P, Anttila A, et al. Performance of colorectal cancer screening in the European Union Member States: data from the second European screening report. Gut. 2018.
40. 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.
41. Spada C, Hassan C, Munoz-Navas M, et al. Second-generation colon capsule endoscopy compared with colonoscopy. Gastrointest Endosc. 2011;74(3):581-9 e1.
42. Van Gossum A, Munoz-Navas M, Fernandez-Urien I, et al. Capsule endoscopy versus colonoscopy for the detection of polyps and cancer. N Engl J Med. 2009;361(3):264-70.

43. Atkin W, Wooldrage K, Parkin DM, et al. Long term effects of once-only flexible sigmoidoscopy screening after 17 years of follow-up: the UK Flexible Sigmoidoscopy Screening randomised controlled trial. Lancet. 2017;389(10076):1299-311.
44. 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.
45. 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.
46. 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.
47. Mandel JS, Church TR, Ederer F, et al. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. J Natl Cancer Inst. 1999;91(5):434-7.
48. Scholefield JH, Moss SM, Mangham CM, et al. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut. 2012;61(7):1036-40.
49. 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.
50. Chiu HM, Chen SL, Yen AM, et al. Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. Cancer. 2015;121(18):3221-9.
51. 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.
52. Lauby-Secretan B, Vilahur N, Bianchini F, et al. International Agency for Research on Cancer Handbook Working G. The IARC Perspective on Colorectal Cancer Screening. N Engl J Med. 2018;378(18):1734-40.
53. Rossi G, Cerquetella M, Pengo G, et al. Immunohistochemical expression of ornithine decarboxylase, diamine oxidase, putrescine, and spermine in normal canine enterocolic mucosa, in chronic colitis, and in colorectal cancer. Biomed Res Int. 2015;2015:172756.
54. Ventura L, Mantellini P, Grazzini G, et al. The impact of immunochemical faecal occult blood testing on colorectal cancer incidence. Dig Liver Dis. 2014;46(1):82-6.
55. Bretthauer M, Kaminski MF, Loberg M, et al. Population-Based Colonoscopy Screening for Colorectal Cancer: A Randomized Clinical Trial. JAMA Intern Med. 2016;176(7):894-902.
56. 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.
57. 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.
58. Laing SS, Bogart A, Chubak J, et al. Psychological distress after a positive fecal occult blood test result among members of an integrated healthcare delivery system. Cancer Epidemiol Biomarkers Prev. 2014;23(1):154-9.
59. Vermeer NC, Snijders HS, Holman FA, et al. Colorectal cancer screening: Systematic review of screen-related morbidity and mortality. Cancer Treat Rev. 2017;54:87-98.
60. van Dam L, Korfage IJ, Kuipers EJ, et al. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? Eur J Cancer. 2013;49(10):2321-30.
61. Lin JS, Piper MA, Perdue LA, et al. Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2016;315(23):2576-94.

62. Kim SY, Kim HS, Park HJ. Adverse events related to colonoscopy: Global trends and future challenges. World J Gastroenterol. 2019;25(2):190-204.
63. Kalager M, Wieszczy P, Lansdorp-Vogelaar I, et al. Overdiagnosis in Colorectal Cancer Screening: Time to Acknowledge a Blind Spot. Gastroenterology. 2018;155(3):592-5.
64. Brouwer NPM, Bos A, Lemmens V, et al. An overview of 25 years of incidence, treatment and outcome of colorectal cancer patients. Int J Cancer. 2018;143(11):2758-66.
65. Schoenberg MH. Physical Activity and Nutrition in Primary and Tertiary Prevention of Colorectal Cancer. Visc Med. 2016;32(3):199-204.
66. 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.
67. 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.
68. Gupta S, Halm EA, Rockey DC, et al. Comparative effectiveness of fecal immunochemical test outreach, colonoscopy outreach, and usual care for boosting colorectal cancer screening among the underserved: a randomized clinical trial. JAMA Intern Med. 2013;173(18):1725-32.
69. Liles EG, Perrin N, Rosales AG, et al. Change to FIT increased CRC screening rates: evaluation of a US screening outreach program. Am J Manag Care. 2012;18(10):588-95.
70. 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.
71. 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.
72. 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.
73. 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.
74. 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.
75. 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.
76. 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.
77. 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.
78. Health Council of the Netherlands. Advisory report A national colorectal cancer screening programme. 2009. [Cited February 12 2019].
79. Zappa M, Castiglione G, Paci E, blood testing: the District of Florence experience. Int J Cancer. 2001;92(1):151-4.
80. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on Cancer. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation of screening outcomes. Endoscopy. 2012;44 Suppl 3:SE49-64.

2

REAL-TIME MONITORING OF RESULTS DURING FIRST YEAR OF DUTCH COLORECTAL CANCER SCREENING PROGRAMME AND OPTIMIZATION BY ALTERING FAECAL IMMUNOCHEMICAL TEST CUT-OFF LEVELS

Esther Toes-Zoutendijk, Monique E. van Leerdam, Evelien Dekker, Frank van Hees, Corine Penning, Iris D. Nagtegaal, Miriam P. van der Meulen, Anneke J. van Vuuren, Ernst J. Kuipers, Johannes M.G. Bonfrer, Katharina Biermann, Maarten G.J. Thomeer, Harriët van Veldhuizen, Sonja Kroep, Marjolein van Ballegooijen, Gerrit A. Meijer, Harry J. de Koning, Manon C.W. Spaander, Iris Lansdorp-Vogelaar, on behalf of the Dutch national colorectal cancer screening working group.

Gastroenterology, 2017

ABSTRACT

Background

After careful pilot studies and planning, the national screening programme for colorectal cancer (CRC), with biennial faecal immunochemical tests (FITs), was initiated in the Netherlands in 2014. A national information system for real-time monitoring was developed to allow for timely evaluation. Data were collected from the first year of this screening programme to determine the importance of planning and monitoring for optimal screening programme performance.

Methods

The national information system of the CRC screening programme kept track of the number of invitations sent in 2014, FIT kits returned, and colonoscopies performed. Age-adjusted rates of participation, the number of positive test results, and positive predictive values (PPVs) for advanced neoplasia were determined weekly, quarterly, and yearly.

Results

In 2014, there were 741,914 persons invited for FIT; of these, 529,056 (71.3%, 95%CI: 71.2-71.4%) participated. A few months into the programme, real-time monitoring showed that rates of participation and positive test results (10.6%, 95%CI: 10.5-10.8%) were higher than predicted and the PPV was lower (42.1%, 95%CI: 41.3-42.9%) than predicted based on pilot studies. To reduce the burden of unnecessary colonoscopies and alleviate colonoscopy capacity, the cut-off level for a positive FIT result was increased from 15 to 47 µg Hb/g faeces halfway through 2014. This adjustment decreased the percentage of positive test results to 6.7% (95%CI: 6.6-6.8%) and increased the PPV to 49.1% (95%CI: 48.3-49.9%). In total, the first year of the Dutch screening programme resulted in the detection of 2,483 cancers and 12,030 advanced adenomas.

Conclusions

Close monitoring of the implementation of the Dutch national CRC screening programme allowed for instant adjustment of the FIT cut-off levels to optimise programme performance.

INTRODUCTION

Colorectal cancer (CRC) is a major health problem.[1] Fortunately, CRC is very suitable for screening and many countries have started CRC screening in the past decade. Choices for screening modality and strategy differ. Worldwide, many countries have implemented faecal occult blood testing (FOBT), in particular by means of faecal immunochemical testing (FIT).[2-4] Various FIT-based initiatives arise based on the growing recognition that an optimal screening method depends on population preference and the availability of resources.[5-9]

In the Netherlands, the screening modality was determined after a period of careful piloting. These pilot studies and subsequent modelling showed that screening by FIT was most acceptable to the Dutch population with a participation rate of up to 60%-62% in the first round, compared to 47%-50% for guaiac-based FOBT (gFOBT), 32% for sigmoidoscopy, 22% for colonoscopy and 34% for computed tomography colonography (CTC). As a result, FIT outperformed the other screening modalities in the detection of CRC per 1000 invitees.[10-14] FIT further allowed for adjustment of the cut-off level enabling a desired balance between true and false positive test results and colonoscopy referral rates to meet colonoscopy resource.[11,15]

Based on these findings, the Dutch government decided to gradually implement a national population-based screening programme based on biennial FIT from age 55 to 75 years at a cut-off level of 15 µg Hb/g faeces. During a 2-year planning period, a national information system was developed for real-time monitoring. Implementation of screening programmes requires careful planning, real-time monitoring and adjustment if needed to achieve the intended impact. Unfortunately, there is limited experience and literature on this process, which is relevant from a clinical as well as a public health perspective.

This article presents the outcomes of the first year of the Dutch CRC screening programme to illustrate the importance of planning and monitoring for optimal screening programme performance.

MATERIALS AND METHODS

The Dutch CRC screening programme

The Dutch CRC screening programme was implemented gradually by age group from 2014 onward, with a projected roll-out period of 5 years, allowing for timely increase of the colonoscopy capacity to ultimately accommodate the target population of 2.2 million invitees annually (Appendix I). The target population for 2014 consisted of all individuals reaching the age of 63, 65, 67, or 75 years in 2014. The oldest age group was included in 2014, because it was their only opportunity to be invited. The age groups around the median age of the programme were selected because these were expected to have the optimal

balance between CRC risk and remaining life-expectancy and experience the highest benefit from screening. Because the programme originally was supposed to start in 2013 and it had been publicly communicated that it would include screening for subjects born in 1938, these individuals also were invited despite having reached the age of 76 years in 2014. The target population received a pre-invitation letter by mail, followed 1 week later by an invitation letter by mail together with a single FIT test (FOB-Gold, Sentinel, Milan, Italy). After 42 days a reminder was sent automatically to nonresponders.

Each invitee was asked to perform a FIT and fill out a reply form including a sample date and return this in a prepaid envelope. Returning the FIT is considered informed consent, in accordance with the Dutch population screening act. The screening programme has been reviewed and approved by the Health Council as part of this act. Participants were informed about the FIT result by mail. If the FIT result equalled or exceeded the cut-off level, the family physician was informed and the participant was invited for a precolonoscopy intake interview in an accredited colonoscopy centre nearby. Participants whose sample was unreliable or not assessable were sent a new test. Individuals who actively deregistered from the programme were labelled as nonparticipants. Individuals who did not respond to the invitation were labelled as nonresponders.

Colonoscopy was the standard diagnostic follow-up test. All colonoscopies were performed by accredited endoscopists who perform at least 300 colonoscopies each year. All detected polyps were to be removed and sent for pathologic review.[16] In case of advanced adenoma (AA) or CRC, the participant was referred for further treatment and surveillance.[17]

Monitoring System

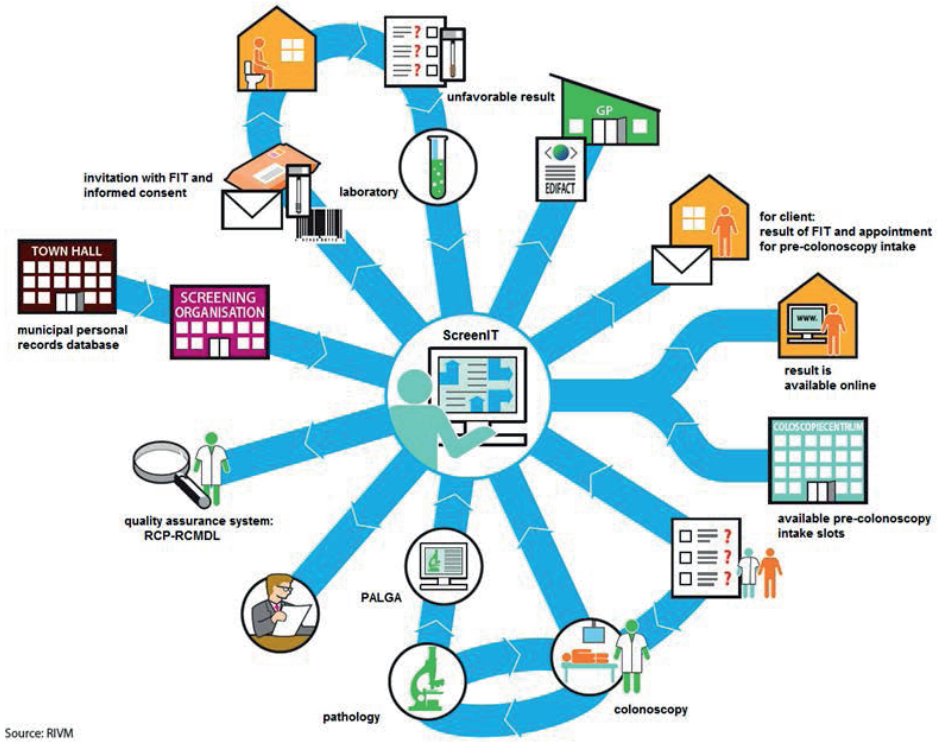
A national information system (ScreenIT, Topicus, Deventer, the Netherlands) was developed to structure the screening process automatically, continuously integrate information from different sources such as endoscopy units and pathology laboratories, and facilitate real-time monitoring (Figure 1). ScreenIT includes personal data from the municipal Personal Records database (personal details of every resident of the Netherlands), FIT results from the laboratories, available pre-colonoscopy intake slots, colonoscopy results from endoscopy centres, and pathology diagnoses from the Dutch national pathology registry (PALGA). Individuals had the right to object to data exchange for scientific research or quality assurance. Those who objected were labelled as nonresponders (n=24).

Screening outcomes from ScreenIT are reported weekly, quarterly, and yearly to the 5 regional screening organisations that are responsible for the execution of the programme.

Programme performance

By using the outcomes of the Dutch pilot studies with FIT (OC Sensor; Eiken Chemical, Tokyo, Japan) as a reference, the programme was designed to accommodate a 60% participation rate with FIT, with a positivity rate of 6.4% at a cut-off level of 15 µg Hb/g faeces. At this cut-

Figure 1: Graphical representation of the workflow in ScreenIT Information System*



Abbreviations: PALGA (pathology database), GP (general practitioner), MDL (gastroenterology), RCP-RCMDL (quality assurance system). Note: abbreviations are based on Dutch descriptions.

*The national information system ScreenIT automatically structures the screening process. It continuously integrates information from different sources, personal data from the municipal Personal Records database, like available pre-colonoscopy intake slots, pathology results from the national pathology registry PALGA, and endoscopy results.

off level the expected positive predictive value (PPV) for CRC and AA combined was 51.6%, and detection rates of CRC were 4.5‰ and of AA were 23.8‰.

Outcomes and Analyses

Data were collected to assess FIT participation rate, positivity rate, participation rate of precolonoscopy intake and diagnostic colonoscopy, PPV for advanced neoplasia, detection rate and false positive rate. Data on the invitees of 2014 were collected until March 31 2015. The FIT participation rate was defined as the number of individuals returning the stool sample divided by the number of individuals invited. The positivity rate was defined as the number of participants with a test result at or above the cut-off level divided by the number of participants with an assessable stool sample. The participation rate for precolonoscopy intake was defined as the number of participants who attended the pre-colonoscopy intake

divided by the number of persons with a positive FIT. The participation rate for colonoscopy was formulated as the number of persons who underwent a colonoscopy divided by the number of persons with a positive FIT. Advanced neoplasia (AN) was considered a relevant abnormality within a CRC screening programme.[18] AN was defined as CRC or any adenoma with histology showing 25% or greater villous component or high-grade dysplasia or adenoma with size 10 mm or larger. The PPV was calculated as the number of persons with AN divided by the number of persons who underwent a colonoscopy. The detection rate was defined as the proportion of individuals with AN detected during colonoscopy per 1,000 screened individuals with an assessable stool sample, also called the true positive rate. The false positive rate was defined as the number of persons without AN detected during colonoscopy divided by the number of screened persons with an assessable stool sample.

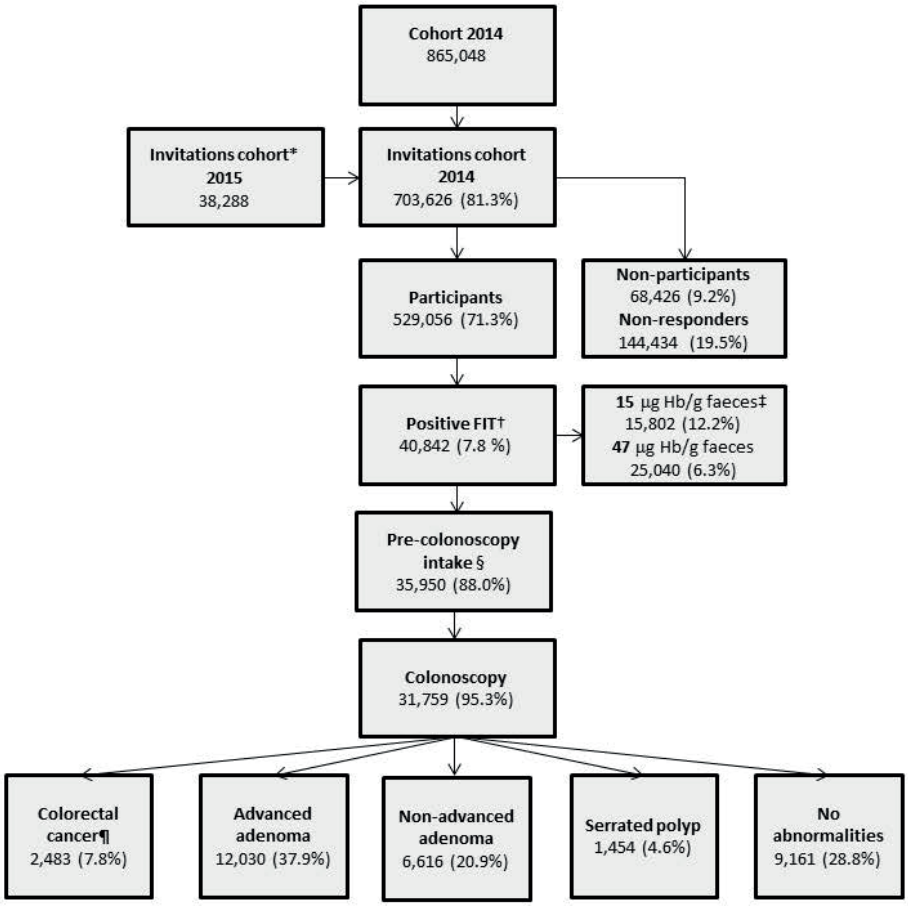
Proportions with 95% CIs were determined by descriptive analyses. Subgroup rates were age-adjusted to the age distribution of the total population invited calculated with a direct standardisation procedure.

RESULTS

Invitation and Participation

The target population for 2014 consisted of 865,048 persons. By the end of the year, 703,626 (81.3%) of those had been invited for screening. Weekly monitoring showed that in some screening regions the entire target population of 2014 had been invited before the end of the year. In these regions, an additional 38,288 persons aged 60 years were invited for screening, resulting in 741,914 invitees in total. Figure 2 shows the flow of individuals through the screening process. A total of 529,056 or 71.3% (95%CI: 71.2-71.4%) of the invitees returned the FIT to the laboratory. Of the 212,858 persons not returning a FIT, 32.1% were classified as nonparticipants and 67.9% were classified as nonresponders (including 24 who objected to data exchange). Of the 529,056 participants, 524,095 (99.1%) had an assessable FIT with consent form. Overall, the test result was positive for 40,842 individuals or 7.8% (95%CI: 7.7-7.9%) however these rates were different for the first half of the year compared with the second half year, as will be discussed in the next section. Of all individuals who tested positive, 35,950 (88.0%) had a precolonoscopy intake interview. In total, 33,313 (92.7%) individuals were advised to undergo colonoscopy. Colonoscopy and pathology data were available for 31,759 (95.3%) of the individuals who were recommended to undergo colonoscopy. Taken together, 77.8% of the participants with a positive FIT had undergone a colonoscopy. Excluding those for whom colonoscopy was not recommended (n=2,637), uptake of colonoscopy was 83.1%.

Figure 2: Flow of individuals through the screening process



Abbreviations: FIT (faecal immunochemical testing)

*As some screening areas had invited the entire target population already before the end of 2014, a number of individuals from the target population of 2015 were already invited in calendar year 2014.

† Including 24 individuals who objected to data exchange, who were also labelled as non-responders.

‡ Of all participants, 99.1% had an assessable FIT.

§ July 2014 the cut-off level for positivity was increased to 47 µg Hb/g faeces.

¶ Preceding the colonoscopy, a pre-colonoscopy intake interview takes places at an accredited screening colonoscopy centre. For 259 participants, no intake report was available in ScreenIT.

** Outcomes are based on the most advanced finding for each individual.

Positivity rate, PPV, and detection rate in the first half of 2014

During the first months of the programme, real-time monitoring detected an age-adjusted positivity rate of 10.6% (95%CI: 10.5-10.8%) at a cut-off level of 15 µg Hb/g faeces. At this cut-off level, the PPV for CRC and AA was 42.1% (95%CI: 41.3-42.9%) and detection rates of CRC and AA were 5.8‰ (95%CI: 5.5-6.1‰) and 30.8‰ (95%CI: 30.1-31.5‰), respectively

Table 1: Participation rates, positivity rates*, positive predictive values (PPVs) †and detection rates‡ for biennial FIT screening in the first year of the Dutch CRC screening programme

		Participation		Positivity		PPV CRC		PPV AA		Detection rates CRC		Detection rates AA	
	Cut-off	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)	n	‰ (95%CI)	n	‰ (95%CI)
All	*Total*	529,056	*71.3 (71.2-71.3)*	40,842	7.8 (7.7-7.9)	2,483	7.8 (7.5-8.1)	12,030	37.9 (37.3-38.4)	2,483	4.7 (4.6-4.9)	12,030	23.0 (22.6-23.4)
Men		256,737	70.3 (70.1-70.4)	24,221	9.5 (9.4-9.6)	1,516	8.0 (7.6-8.4)	7,761	41.0 (40.3-41.7)	1,516	6.0 (5.7-6.3)	7,761	30.5 (29.9-31.2)
Women		272,319	72.3 (72.2-72.5)	16,621	6.2 (6.1-6.3)	967	7.5 (7.1-8.0)	4,269	33.3 (32.5-34.1)	967	3.6 (3.4-3.8)	4,269	15.8 (15.4-16.3)
60		26,622	69.5 (69.1-70.0)	1,310	5.0 (4.7-5.3)	38	4.3 (3.2-5.9)	332	37.8 (34.6-41.0)	38	1.5 (1.1-2.0)	332	12.7 (11.4-14.1)
63		89,420	74.2 (73.9-74.4)	4,842	5.5 (5.3-5.6)	273	7.3 (6.5-8.1)	1,475	39.2 (37.7-40.8)	273	3.1 (2.7-3.5)	1,475	16.7 (15.8-17.5)
65		116,998	74.3 (74.1-74.5)	7,906	6.8 (6.7-7.0)	421	6.7 (6.1-7.3)	2,503	39.6 (38.4-40.8)	421	3.6 (3.3-4.0)	2,503	21.6 (20.8-22.4)
67		141,682	74.5 (74.3-74.7)	9,593	6.8 (6.7-7.0)	633	8.2 (7.6-8.9)	3,110	40.4 (39.3-41.5)	633	4.5 (4.2-4.9)	3,110	22.1 (21.4-22.9)
75		79,095	67.1 (66.8-67.3)	7,789	9.9 (9.7-10.2)	538	9.2 (8.5-9.9)	2,188	37.3 (36.0-38.5)	538	6.9 (6.3-7.5)	2,188	28.0 (26.8-29.1)
76		75,239	64.1 (63.8-64.4)	9,402	126 (12.4-12.9)	580	8.0 (7.4-8.7)	2,422	33.5 (32.4-34.6)	580	7.8 (7.2-8.5)	2,422	32.6 (31.3-33.9)
All	*15 µg*	130,457		15,802	12.2 (12.0-12.4)	911	7.2 (6.8-7.7)	4,319	34.3 (33.4-35.1)	911	7.0 (6.6-7.5)	4,319	33.4 (32.4-34.4)
age-adjusted§				*15,802*	*10.6 (10.5-10.8)*	*911*	*6.8 (6.4-7.2)*	*4,319*	*35.3 (34.6-36.1)*	*911*	*5.8 (5.5-6.1)*	*4,319*	*30.8 (30.1-31.5)*
60		-		-	-	-		-		-	-	-	-
63		2,709		207	7.7 (6.7-8.8)	8	4.6 (2.3-8.9)	63	36.0 (29.2-43.4)	8	3.0 (1.5-5.9)	63	23.4 (18.3-29.8)
65		28,812		2,598	9.1 (8.8-9.4)	123	5.6 (4.7-6.6)	816	37.0 (35.0-39.0)	123	4.3 (3.6-5.10	816	28.5 (26.7-30.5)
67		16,726		1,753	10.5 (10.1-11.0)	102	7.0 (5.8-8.5)	534	36.9 (34.5-39.4)	102	6.1 (5.1-7.4)	534	32.1 (29.5-34.9)
75		26,446		3,454	13.2 (12.8-13.6)	217	8.0 (7.1-9.1)	938	34.8 (33.0-36.6)	217	8.3 (7.2-9.4)	938	35.8 (33.6-38.1)
76		55,764		7,790	14.1 (13.8-14.4)	461	7.6 (6.9-8.3)	1,968	32.3 (31.2-33.5)	461	8.4 (7.6-9.2)	1,968	35.7 (34.2-37.3)
All	*47 µg*	398,599		25,040	6.3 (6.3-6.4)	1,572	8.2 (7.8-8.6)	7,711	40.3 (39.6-41.0)	1,572	4.0 (3.8-4.2)	7,711	19.5 (19.1-20.0)
age-adjusted§				*23,730*	*6.7 (6.6-6.8)*	*1,534*	*8.9 (8.4-9.3)*	*7,379*	*40.2 (39.5-41.0)*	*1,534*	*4.4 (4.2-4.7)*	*7,379*	*20.6 (20.0-21.2)*
60		26,622		1,310	5.0 (4.7-5.3)	38	4.3 (3.2-5.9)	332	37.8 (34.6-41.0)	38	1.5 (1.1-2.0)	332	12.7 (11.4-14.1)
63		86,711		4,635	5.4 (5.2-5.5)	265	7.4 (6.6-8.3)	1,412	39.4 (37.8-41.0)	265	3.1 (2.7-3.5)	1,412	16.4 (15.6-17.3)
65		88,186		5,308	6.1 (5.9-6.2)	298	7.2 (6.5-8.0)	1,687	41.0 (39.5-42.5)	298	3.4 (3.1-3.8)	1,687	19.3 (18.4-20.3)

Table 1: Participation rates, positivity rates*, positive predictive values (PPVs) †and detection rates‡ for biennial FIT screening in the first year of the Dutch CRC screening programme (continued)

	Participation			Positivity		PPV CRC		PPV AA		Detection rates CRC		Detection rates AA	
	Cut-off	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)	n	‰ (95%CI)	n	‰ (95%CI)
67		124,956		7,840	6.3 (6.2-6.5)	531	8.5 (7.8-9.2)	2,576	41.2 (40.0-42.4)	531	4.3 (3.9-4.7)	2,576	20.8 (20.0-21.6)
75		52,649		4,335	8.3 (8.1-8.6)	321	10.1 (9.1-11.2)	1,250	39.4 (37.7-41.1)	321	6.2 (5.5-6.9)	1,250	24.0 (22.7-25.4)
76		19,475		1,612	8.4 (8.0-8.8)	119	10.4 (8.7-12.3)	454	39.7 (36.9-42.5)	119	6.2 (5.2-7.4)	454	23.6 (21.6-25.9)

Abbreviations: FIT (faecal immunochemical testing), CRC (colorectal cancer), AA (advanced adenomas), PPV (positive predictive value)

* Positivity rate was defined as the number of participants with an unfavourable test result (above the cut-off level) divided by the number of participants with assessable stool sample. Numbers of participants with assessable stool sample are not shown in the table.

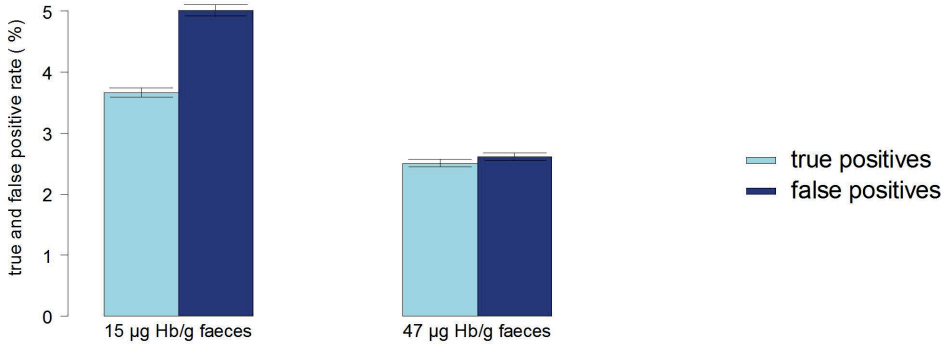
† PPV was calculated as the number of persons with CRC or AA divided by the number of persons who underwent colonoscopy. Numbers of positive individuals attending colonoscopy are not shown.

‡ Detection rate was defined as the proportion of persons with CRC or AA detected during colonoscopy per 1,000 screened persons with assessable stool sample. Numbers of participants with assessable stool sample are not shown.

§ The age-adjusted rates are calculated with the exclusion of the 60-year-olds screened in the second half of 2014 (with the cut-off level 47 µg Hb/g faeces).

(Table 1). Both positivity (10.6% vs 6.4%) and detection rates of CRC (5.8‰ vs 4.5‰) and AA (30.8‰ vs 23.8‰) in the first half of the year were higher than the expected programme performance. However, the PPV for CRC and AA was lower than expected (42.1% vs 51.6%), with relatively more individuals having a false positive result. The false-positive rate was 5.0% (95%CI: 4.9-5.1%), resulting in a higher burden of colonoscopy for both participating individuals and the programme. In addition, the participation rate also was higher than expected (71% vs 60%). Consequently, the demand for colonoscopies exceeded the capacity leading to a prolonged waiting period. There was no excess colonoscopy capacity in the Netherlands as a whole, so a further increase in colonoscopy capacity for the national programme was not possible in the short term. Because the programme was not performing according to the predefined quality indicators (i.e., positivity rate of 6.4; PPV of 51.6%; and follow-up colonoscopy within 3 weeks after a positive FIT), an immediate decision had to be made to improve the programme. A decision analysis was performed comparing 3 different methods to decrease colonoscopy demand in 2014: increase cut-off level, postpone screening in selected age groups, and forego screening in older age groups. This analysis showed that increasing the cut-off level not only resulted in the lowest decrease in CRC deaths prevented, but also resulted in a balance between harms and benefits of screening in accordance with the aims at programme start.[19] In consultation with all stakeholders, the Dutch National Institute for Public Health the Environment (RIVM) decided to increase the cut-off level because this was the most efficient way to optimise programme performance. Therefore, the cut-off level for referral for colonoscopy was increased to 47 µg Hb/g faeces in July 2014.

Figure 3: Comparison of the balance between true and false positives by the two cut-off levels



Abbreviations: FIT (faecal immunochemical testing), Hb (haemoglobin)

† True positive rate was defined as the number of persons with CRC or AA detected during colonoscopy divided by the number of screened persons with assessable stool sample.

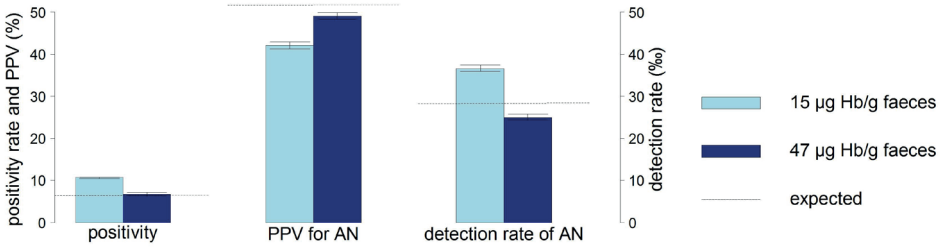
‡ True negative rate was defined as the number of persons without CRC or AA detected during colonoscopy divided by the number of screened persons with assessable stool sample.

§Rates are presented as age-adjusted rates, calculated with the exclusion of the 60-year-olds screened in the second half of 2014 (with the cut-off level 47 µg Hb/g faeces).

Positivity rate, PPV, and detection rate in the second half of 2014

The adjustment of the programme resulted in an age-adjusted positivity rate of 6.7% (95%CI: 6.6-6.8%) at a cut-off level of 47 µg Hb/g faeces. At this cut-off level the age-adjusted PPV for CRC and AA was 49.1% (95%CI: 48.3-49.9%). The age-adjusted detection rates at 47 µg Hb/g faeces of CRC and AA were 4.4‰ (95%CI: 4.2-4.7‰) and 20.6‰ (95%CI: 20.0-21.2‰), respectively. The false positive rate was 2.6% (95%CI: 2.6-2.7%). Increasing the cut-off level halfway through 2014 decreased the demand for colonoscopies by 37% and the number of false positive results by 48%, although CRC and AA detection rates decreased to a lesser extent by only 23% and 33%, respectively (Figure 3 and 4). If we would have applied a cut-off level of 47 µg Hb/g faeces during the first half year, 6433 (40.7%) fewer participants would have tested positive. This would have led to failure to detect 132 of 911 (14.5%) CRCs and 1351 of 4319 (31.3%) AAs (Appendix 2).

Figure 4: Comparison of age-adjusted† positivity rates‡, positive predictive value and detection rates§ by the two cut-off levels



Abbreviations: PPV (positive predictive value), AN (advanced neoplasia), Hb (haemoglobin)

* Comparing the two cut-off levels, the positivity rates, PPVs and detection rates were significantly different (p<0.05)

† All rates are presented as age-adjusted rates, with the exclusion of the 60-year-olds screened in the second half of 2014 (with the cut-off level 47 µg Hb/g faeces).

‡ Positivity rate was defined as the number of participants with an unfavourable test result (above the cut-off level) divided by the number of participants with assessable stool sample.

§ PPV was calculated as the number of persons with CRC or AA divided by the number of persons who underwent colonoscopy.

¶ Detection rate was defined as the proportion of persons with CRC or AA detected during colonoscopy per 1,000 screened persons with assessable stool sample.

DISCUSSION

These data show the additional value of real-time monitoring to successfully implement a national screening programme. A few months into the programme real-time monitoring showed a higher positivity rate and a lower PPV than expected. This resulted in a higher number of false positive test results, leading to unnecessary diagnostic colonoscopies with

associated risks. In July 2014, the programme was adjusted resulting in a lower positivity rate and fewer false positive results, which was more in line with expectations. Despite this adjustment, the entire target population of 2014 could not be invited within that calendar year owing to high participation and high referral rates for colonoscopy in the first half of the year, together leading to higher colonoscopy demand than capacity. As a whole, the first year of the national CRC screening programme resulted in a high number of participants and the detection of 2483 cancers and 12,030 advanced adenomas.

The participation rate exceeded the level of 65%, put forward as the desirable level of participation in the European Union guidelines for quality assurance in CRC screening.[18] It is remarkable that some countries with similar participation rates for population-based breast cancer screening programmes as the Netherlands, such as England and France, have lower participation rates for CRC screening.[20-22] Potential explanations for the high participation rate are the choice of screening modality, the efficient organisational structure of the programme (e.g., invitation-based, pre-invitation letter, and enclosing the FIT within the invitation set), a detailed information leaflet, and the fact that the FIT is free of charge.[23]

The higher-than-expected positivity rate in the first year can be explained by several factors. First, the age distribution of the first half of the year was skewed toward older ages, and older age is related to a higher positivity rate.[24] Second, the FIT (FOB-Gold) in the national programme was of a different brand than the FIT in the pilot studies (OC-Sensor). The choice of FIT in the programme was the result of a public tender required for all governmental purchases exceeding a certain monetary value. Previous studies have shown that FOB-Gold has higher positivity rates than OC-Sensor.[25,26] However, these studies did not standardise a cut-off value for a positive test result in µg Hb/g faeces, which was performed in the Dutch programme. Equal performance of both tests later was assessed in a confirmative trial.[27] Counterintuitively, positivity rates for both tests were higher in this fourth round of screening than in the third round, indicating a change in test performance between the 2 study rounds. The explanation may be that the manufacturers recently adapted the composition of the buffer fluid in the collection device to improve sample stability. Nevertheless, because selecting a new screening test is always subject to public tender, in the future we might first pilot a newly chosen test.

Participation in diagnostic colonoscopy was short of the minimally acceptable level of 85% and was lower than in the pilot studies and other FOBT-based screening programmes.[3,10,14,28,29] One explanation may be that not all colonoscopy results were integrated in ScreenIT because some individuals may have had a colonoscopy in centres outside the screening programme. Another potential explanation is that colonoscopy costs are considered standard medical care and therefore are covered by health care insurance. Because all Dutch citizens have a yearly obligatory deductible excess of 360 Euro participants who did not have previous medical costs in the calendar year were obliged to pay part of the colonoscopy costs. One last explanation might be that individuals refrained from colonoscopy or that colonoscopy was

not considered appropriate because of comorbidities. This also was reflected in the higher participation rate when excluding those for whom colonoscopy was not recommended. This may have had a relatively large impact because older age groups were disproportionally present in the 2014 target population.

The success of the Dutch CRC screening programme can in large part be attributed to its coordinated preparation and implementation, including piloting and monitoring. The piloting phase allowed for an evidence-based choice of the screening test, cut-off level, interval, and age range for the Dutch setting.[11-14] Real-time monitoring showed at an early stage that the programme performed differently than intended (e.g., higher positivity rate and a lower PPV) and therefore adjustments to the programme could be made immediately. Consequently, the programme now performs in line with expectations and recommendation of the Dutch Health Council.[3,11-14] The decision to adjust the cut-off level of the test was based on a thorough decision analysis, which has shown that the adjustment will lead to similar long-term effectiveness as the intended programme; close monitoring of programme performance should be continued for 2 reasons.[19] First, given the gradual implementation of the programme, the new cut-off level could be based on only a selected number of age groups. Second, results were all based on the first screening round. Data on participation rate, positivity rate, PPV, and the detection rate of subsequent rounds will be of interest to evaluate whether the chosen cut-off level will continue to provide the expected results. In the end, the interval cancer rate will be the most important performance parameter to monitor the adjustment of the programme.

One may argue that it is unethical to increase the cut-off level of a screening test after implementation of a screening programme, because an increase in the cut-off value would decrease detection of advanced neoplasia, and screening should not be implemented unless sufficient resources for follow-up evaluation and treatment have been secured.[30] However, in our opinion this consideration is not applicable to the Dutch programme. The increase of the cut-off value has not been made in the consideration of colonoscopy capacity (alone), but rather because the chosen cut-off value did not result in the intended balance of harms and benefits of the screening programme as recommended by the Dutch Health Council.[3] The adjustments of the cut-off value halfway through the year was necessary to ensure that the programme again met this intended performance.

Three limitations of monitoring the Dutch CRC screening programme are noteworthy. First, at this point in time, there is a delay of reliable information on the stage and localisation distribution of the detected CRCs and the occurrence of adverse events because not all data sources have been completely linked to ScreenIT. It is expected that this will occur within the next 1-2 years. Second, only colonoscopies performed in accredited colonoscopy centres within the programme are reported in ScreenIT. A national colonoscopy database will be set up in 2016 and potential linkage with that database will ensure complete catchment of all

colonoscopies by using the unique personal identifier. Finally, as already mentioned, current results concern only individuals age 60 years or older.

Notwithstanding these limitations, the design of the Dutch CRC screening programme may serve as a best practice for many screening initiatives currently being organised worldwide.[2] The implementation of the Dutch programme illustrates that even in evidence-based, well-planned programmes, programme performance can deviate from planning. Therefore, real-time monitoring systems are indispensable in ensuring quality in all aspects of screening programmes. A considerable amount of literature has been published on reports of organised screening programmes.[28,29] However, to our knowledge no other data on real-time short and long cycle monitoring of screening programmes is available. The results of the first year of the Dutch programme may serve as an example to show that real-time monitoring is different from a retrospective monitoring system. By using real-time monitoring, adjustments can be made instantaneously to obtain optimal programme performance and ensure a good balance between harms and benefits of the programme. The way monitoring is performed may differ throughout the world, to best fit specific conditions. But even in settings with opportunistic screening such as in the United States, monitoring systems can be put in place on local or institutional levels. The Kaiser Permanente Northern California organised CRC screening programme is an excellent example of setting up large organised programmes to document the entire screening process with the purpose of monitoring and quality assurance.[31]

At this phase of implementation, it was decided to stick to the original programme as much as possible and not experiment with alternative programme designs. Therefore, the cut-off level was increased the same way across the board, i.e. for all ages and screening rounds. It also was decided to stick to the re-screening interval of 2 years, even if the target population had not been completely invited yet. Once the programme has been fully implemented and established, future research should be performed to continually optimise the CRC screening programme. Looking at the impact of increasing the cut-off level, there are indications that the higher cut-off level has led to a greater reduction in screen-detected neoplasia in the older age groups compared with the younger age groups. Therefore, a differential cut-off level by age might lead to a more (cost-) effective screening programme. Decision analyses can be performed comparing such strategies to identify the optimal screening strategy for the current situation. Other examples of possible new strategies are applying different cut-off levels for the first and subsequent screening rounds or allocating different intervals based on the haemoglobin level of the previous screening round.

In conclusion, the Dutch national CRC screening programme was implemented successfully with high participation and yield. Real-time monitoring allowed for instant adjustment of the programme when it substantially differed from expected. Optimising the programme resulted in a programme that more closely meets expectations, with a better balance between true and false positive results.

REFERENCES

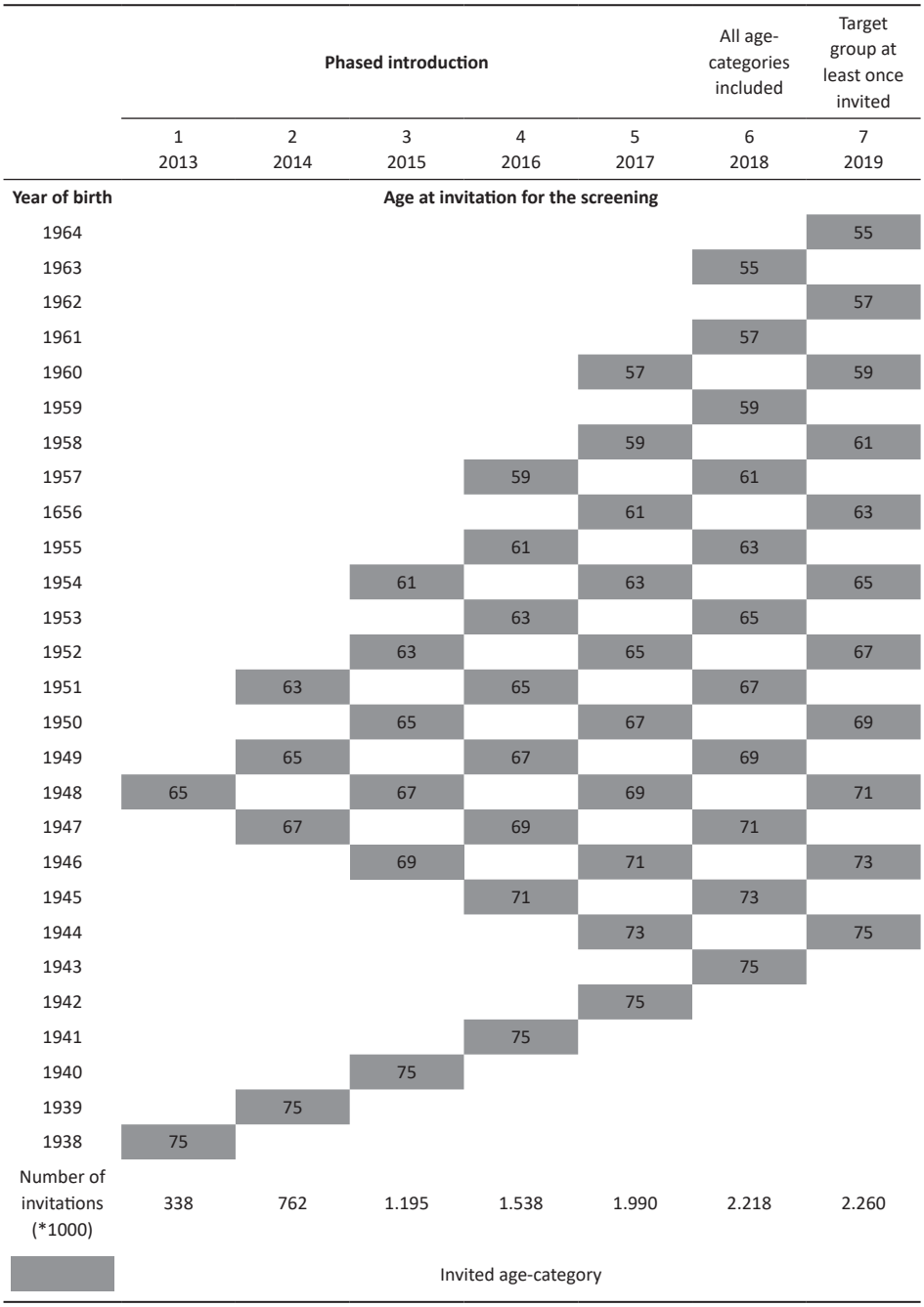
1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin 2015;65(2):87-108.
2. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015; 64(10):1637-49
3. Council of the European Union. Council recommendation of 2 december 2003 on cancer screening (2003/878/EC). Official Journal of the European Union 2003;327:34-38. Luxembourg, Luxembourg: Council of the European Union. http:// eur-lex. europa.eu/LexUriServ/LexUriServ.do?uri-OJ:L:2003:327:0034:0038:EN:PDF. [Cited July 29 2015.
4. AIHW 2015. National Bowel Cancer Screening Programme: monitoring report 2013–14 Cancer series no. 94. Cat. no. CAN 92. Canberra: AIHW. http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129551419. [Cited August 14 2015].
5. Canadian Task Force on Preventive Health Care. Recommendations on screening for colorectal cancer in primary care. Canadian Medical Association Journal 2016; 188 (5), 340-348.
6. van der Steen A, Knudsen AB, van Hees F, et al. Optimal colorectal cancer screening in states' low-income, uninsured populations-the case of South Carolina. Health Serv Res 2015;50(3):768-89.
7. Liles EG, Perrin N, Rosales AG, et al. Change to FIT increased CRC screening rates: evaluation of a US screening outreach programme. Am J Manag Care 2012;18(10):588-95.
8. Gupta S, Halm EA, Rockey DC, et al. Comparative effectiveness of faecal immunochemical test outreach, colonoscopy outreach, and usual care for boosting colorectal cancer screening among the underserved: a randomized clinical trial. JAMA Intern Med 2013;173(18):1725-32.
9. Sung JJY, Ng SC, Chan FKL, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. Gut 2015; 64(1), 121-132.
10. 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.
11. 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.
12. 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.
13. 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.
14. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical faecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;135(1):82-90.
15. Wilschut JA, Habbema JD, van Leerdam ME, et al. Faecal occult blood testing when colonoscopy capacity is limited. J Natl Cancer Inst 2011;103(23):1741-51.
16. Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 2007;29(1):19-24.
17. Dekker E, van Leerdam ME, Hazewinkel Y, et al. Nederlandse Richtlijn Coloscopie Surveillance. Nederlandse Vereniging

van Maag-, Darm- en Leverartsen in samenwerking met Nederlandse Vereniging van Pathologie. 2013. http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf [Cited July 30 2015].

18. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on Cancer. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation of screening outcomes. Endoscopy 2012;44 Suppl 3:SE49-64.
19. 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
20. Giordano L, von Karsa L, Tomatis M, et al. Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen 2012;19 Suppl 1:72-82.
21. Klabunde C, Blom J, Bulliard JL, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. J Med Screen 2015;22(3):119-26.
22. Viguier J, Morere JF, Touboul C, et al. Cancer screening in France: third edition of the EDIFICE survey. Eur J Cancer Prev 2015;24 Suppl:S68-72.
23. 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
24. Symonds EL, Osborne JM, Cole SR, et al. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. J Med Screen 2015; 22(4):187-93.
25. Faivre J, Dancourt V, Denis B, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer 2012;48(16):2969-76.
26. Zubero MB, Arana-Arri E, Pijoan JI, et al. Population-based colorectal cancer screening: comparison of two faecal occult blood test. Front Pharmacol 2014;4:175.
27. Grobbee EJ, van der Vlught M, Stroobants AJ, et al. A randomised comparison of two faecal immunochemical tests in population-based colorectal cancer screening. GUT 2017;66(11):1975-1982.
28. Moss S, Mathews C, Day TJ, et al. Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results form a pilot study within the national screening programmeme in England. GUT 2017;66(9):1631-1644
29. Denis B, Gendre I, Perrin P. Participation in four rounds of a French colorectal cancer screening programme with guaiac faecal occult blood test: a population-based open cohort study. J Med Screen 2015;22(2):76-82.
30. Wilson, James Maxwell Glover, and Gunnar Jungner. “Principles and practice of screening for disease.” World Health Organization, Public Health Paper. 1968; 34.
31. Levin TR, Jamieson L, Burley DA, et al. Organized colorectal cancer screening in integrated health care systems. Epidemiol Rev. 2011; 33:101-10.

APPENDIX I

Figure 1: Original scenario for the gradual implementation of the Dutch national colorectal cancer screening programme



Original scenario for the gradual implementation by age groups to ultimately accommodate the target population of 2.2 million invitees (Source: RIVM).

APPENDIX II

Table 1: Calculated positivity rate* and positive predictive value (PPV) † for biennial FIT screening in the first year of the Dutch CRC screening programme at a cut-off level of 47 µg Hb/g faeces

		First half year 2014				Second half year 2014	
Cut-off level		*original 15 µg*		*15 µg ->47 µg*		*47 µg*	
	Positivity	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)
All		15,802	12.2 (12.0-12.4)	9,369	7.2 (7.1-7.4)	25,040	6.3 (6.3-6.4)
age-adjusted§		*15,802*	*10.6 (10.5-10.8)*	*9,369*	*6.4 (6.3-6.4)*	*23,730*	*6.7 (6.6-6.8)*
60		-	-	-	-	1,310	5.0 (4.7-5.3)
63		207	7.7 (6.7-8.8)	125	4.6 (3.9-5.5)	4,635	5.4 (5.2-5.5)
65		2,598	9.1 (8.8-9.4)	1,533	5.4 (5.1-5.6)	5,308	6.1 (5.9-6.2)
67		1,753	10.5 (10.1-11.0)	1,056	6.3 (6.0-6.7)	7,840	6.3 (6.2-6.5)
75		3,454	13.2 (12.8-13.6)	2,066	7.9 (7.6-8.2)	4,335	8.3 (8.1-8.6)
76		7,790	14.1 (13.8-14.4)	4,589	8.3 (8.1-8.6)	1,612	8.4 (8.0-8.8)
	PPV CRC	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)
All		911	7.2 (6.8-7.7)	779	10.4 (9.7-11.1)	1,572	8.2 (7.8-8.6)
age-adjusted§		*911*	*6.8 (6.4-7.2)*	*779*	*9.7 (9.2-10.2)*	*1,534*	*8.7 (8.3-9.2)*
60		-	-	-	-	38	4.3 (3.2-5.9)
63		8	4.6 (2.3-8.9)	8	7.3 (3.7-14.0)	265	7.4 (6.6-8.3)
65		123	5.6 (4.7-6.6)	109	8.3 (6.9-9.9)	298	7.2 (6.5-8.0)
67		102	7.0 (5.8-8.5)	87	10.0 (8.2-12.2)	531	8.5 (7.8-9.2)
75		217	8.0 (7.1-9.1)	183	11.3 (9.8-12.9)	321	10.1 (9.1-11.2)
76		461	7.6 (6.9-8.3)	392	10.9 (10.0-12.0)	119	10.4 (8.7-12.3)
	PPV AA	n	% (95%CI)	n	% (95%CI)	n	% (95%CI)
All		4,319	34.3 (33.4-35.1)	2,968	39.6 (38.5-40.7)	7,711	40.3 (39.6-41.0)
age-adjusted§		*4,319*	*35.3 (34.6-36.1)*	*2,968*	*40.9 (40.0-41.7)*	*7,379*	*40.3 (39.4-41.1)*
60		-	-	-	-	332	37.8 (34.6-41.0)
63		63	36.0 (29.2-43.4)	42	38.5 (29.9-48.0)	1,412	39.4 (37.8-41.0)
65		816	37.0 (35.0-39.0)	558	42.5 (39.8-45.2)	1,687	41.0 (39.5-42.5)
67		534	36.9 (34.5-39.4)	381	43.7 (40.5-47.1)	2,576	41.2 (40.0-42.4)
75		938	34.8 (33.0-36.6)	653	40.2 (37.8-42.6)	1,250	39.4 (37.7-41.1)
76		1,968	32.3 (31.2-33.5)	1,334	37.3 (35.7-38.8)	454	39.7 (36.9-42.5)

Abbreviations: FIT (faecal immunochemical testing), CRC (colorectal cancer), AA (advanced adenomas), PPV (positive predictive value)

* Positivity rate was defined as the number of participants with an unfavorable test result (above the cut-off level) divided by the number of participants with assessable stool sample. In this table, FITs was considered positive at a cut-off level of 47 µg Hb/g faeces.

† PPV was calculated as the number of persons with CRC or AA divided by the number of persons who underwent colonoscopy. Numbers of positive individuals attending colonoscopy are not shown.

§ The age-adjusted rates are calculated with the exclusion of the 60-year-olds screened in the second half of 2014.

3

STAGE DISTRIBUTION OF SCREEN-DETECTED COLORECTAL CANCERS IN THE NETHERLANDS

Esther Toes-Zoutendijk, Arthur I. Kooyker, Marloes A. Elferink, Manon C.W. Spaander, Evelien Dekker, Harry J. de Koning, Valery E. Lemmens, Monique E. van Leerdam, Iris Lansdorp-Vogelaar, on behalf of LECO working group.

GUT, 2018

We read with interest the results of the pilot study in England, which was performed to establish the acceptability and diagnostic performance of screening with the Faecal Immunochemical Test (FIT) over the guaiac Faecal Occult Blood Test (gFOBT).[1] When comparing gFOBT to FIT, the uptake increased from 59.4% to 66.4%, positivity rate increased from 1.7% to 7.8% (at a cut-off level of 20 µg Hb/g faeces) and colorectal cancer (CRC) detection doubled. Moreover, this report showed that also with FIT cut-off levels above 20 µg Hb/g faeces, improved clinical outcomes can be achieved over gFOBT. However, for a screening programme to be effective it is a prerequisite to detect cancers in an early stage. Thus far, information on stage distribution of screen-detected CRCs in a population-based FIT screening programme is lacking.

In our study, we collected data on all CRCs detected in patients aged 60 to 75 years in the Netherlands in 2015 through the Netherlands Cancer Registry. In total, 9,437 CRCs were diagnosed in 9,301 patients: 3,579 (38.5%) patients were diagnosed after a positive FIT in the CRC screening programme (screen-detected), 4,506 (48.4%) patients were diagnosed due to symptoms (symptom-detected) and 1,216 (13.1%) patients had another mode of detection. Among all patients, 2,679 (31.3%) CRCs were detected at stage I, 1,827 (21.2%)

Table 1: Screen- and symptom-detected colorectal cancers

			Screen-detected (N = 3,521)	**Symptom-detected** (N = 4,455)	
			n (%)	n (%)	*p* value
Stage*	I		1,697 (48.2%)	743 (16.7%)	
	II		652 (18.5%)	1,028 (23.1%)	
	III		957 (27.2%)	1,541 (34.6%)	
	IV		215 (6.1%)	1,143 (25.7%)	<0.001
Location		Right	944 (26.6%)	1,569 (35.4%)	
	Caecum		280 (7.8%)	588 (13.0%)	
	Ascending colon		268 (7.5%)	494 (11.0%)	
	Hepatic flexure of colon		112 (3.1%)	166 (3.7%)	
	Transverse colon		183 (5.1%)	203 (4.5%)	
	Splenic flexure		101 (2.8%)	118 (2.6%)	
		Left	1,628 (45.9%)	1,393 (31.4%)	
	Descending colon		176 (4.9%)	167 (3.7%)	
	Sigmoid		1,422 (39.7%)	1,213 (26.9%)	
	Rectosigmoid		30 (0.8%)	13 (0.3%)	
		Rectum	973 (27.4%)	1,476 (33.3%)	
	Rectum		973 (27.2%)	1,476 (32.8%)	<0.001
	Colon overlapping		11 (0.3%)	24 (0.5%)	
	Colon unspecified		23 (0.6%)	44 (1.0%)	

* Stage distribution of screen-detected and symptom-detected CRCs was unknown of 109 patients.

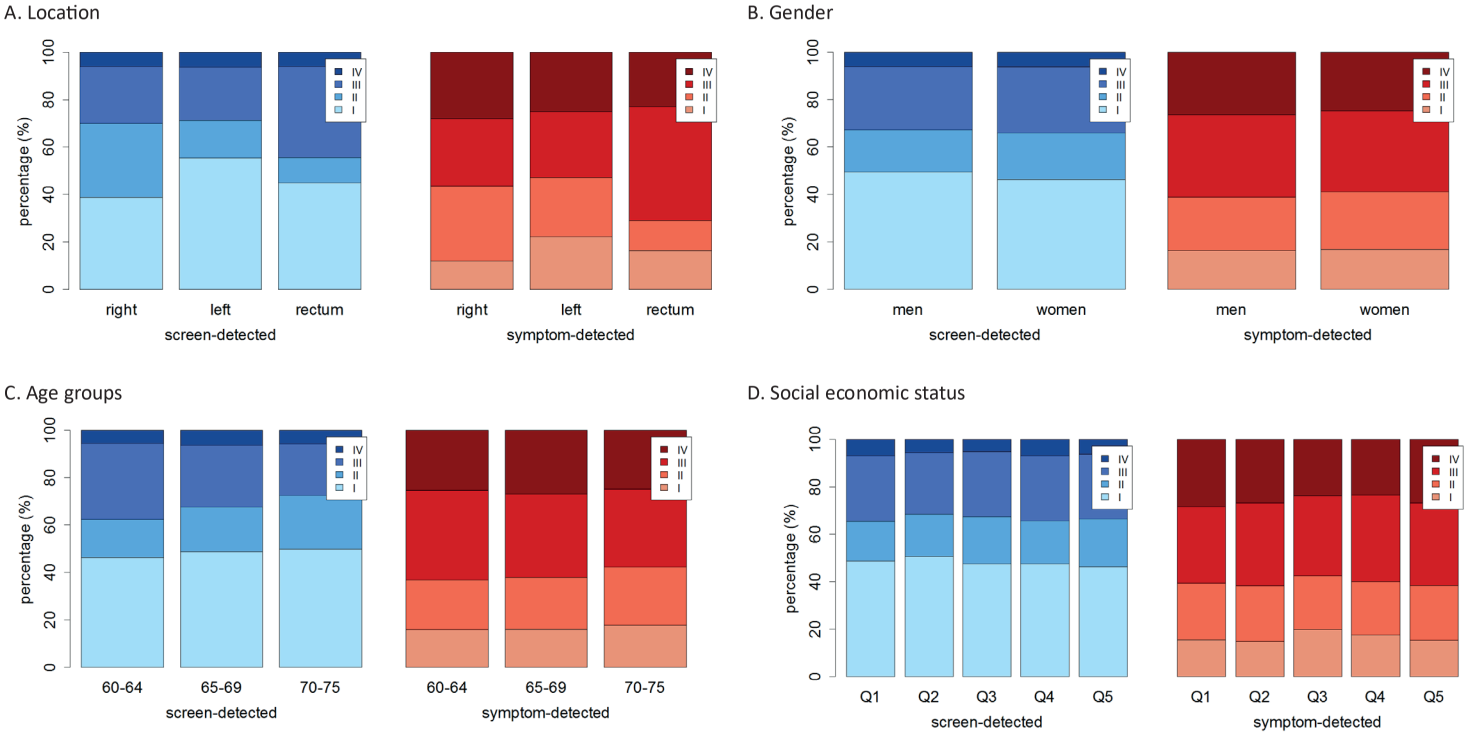
CRCs at stage II, 2,630 (30.6%) CRCs at stage III and 1,469 (17.1%) CRCs at stage IV. CRCs stage was unknown for 696 (7.5%) CRCs. Comparison of only screen-detected and symptom-detected CRCs showed that 2,349 (66.7%) screen-detected CRCs were detected at an early stage (stages I and II), which was higher than the 1,771 (39.8%) symptom-detected CRCs (p <0.001, table 1). Screen-detected cancers were more often diagnosed in the left colon (45.9%) than symptom-detected cancers (31.4%, p <0.001). Figure 1 presents the comparison of screen-detected and symptom-detected CRCs by stage distribution for different subgroups (location, gender, age and social economic status). As shown in this Figure, in all subgroups the screen-detected CRCs had a more favourable stage distribution than the symptom-detected CRCs. This indicates that there is no difference in impact of FIT-based screening on stage distribution among different subgroups.

The more favourable stage distribution of screen-detected CRCs may have two main explanations. The first is simply the earlier detection of cancers by FIT screening, which is the aim. However, it may also be due to overdiagnosis of indolent disease. In this case, screening would simply result in more detection of early-stage disease, without actually reducing the amount of advanced CRCs. Therefore, a more favourable stage distribution of screen-detected versus symptomatic CRCs in itself gives no definitive evidence for future mortality reduction.[2,3]

On the other hand, stage distribution of the screen-detected CRCs is in line with that observed in randomized controlled trials, which eventually demonstrated a reduction in CRC mortality.[4-6] It therefore is expected that the Dutch FIT-based screening programme may also decrease CRC mortality. The observed stage distribution is also in line with other FOBT programmes: the UK gFOBT-based screening programme and the FIT-based screening programme of the Basque country.[7,8] In the British sigmoidoscopy trial, the proportion of CRCs detected at an early stage (74%) was slightly higher, but in line with our FIT programme.[9]

In conclusion, screen-detected CRCs in a FIT-based screening programme are detected more often at an early disease stage than symptomatic cancers. As survival rates improve if cancers are detected at an early stage, the first prerequisite for mortality reduction as a result of FIT-based screening is met.

Figure 1: Stage distribution of screen-detected and symptom-detected colorectal cancers by subgroups



Abbreviations: Q (Quintiles, with quintile 1 having the highest SES score and quintile 5 having the lowest SES score).

REFERENCES

1. Moss S, Mathews C, Day T J, et al. Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results form a pilot study within the national screening programme in England. GUT 2017; 66:1631-1644.
2. Lannin DR, Wang S. Are Small Breast Cancers Good because They Are Small or Small because They Are Good? N Engl J Med 2017;376:2286-91.
3. Murphy CC, Lund JL, Sandler RS. Young-Onset Colorectal Cancer: Earlier Diagnoses or Increasing Disease Burden? Gastroenterology 2017;152:1809-12 e3.
4. 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;328:1365-71.
5. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;369:1106-14.
6. Atkin W, Wooldrage K, Maxwell Parkin D, et al. Long-term effects of once-only flexible sigmoidoscopy screening after 17 years of follow-up: the UK Flexible Sigmoidoscopy Screening randomized controlled trial. Lancet 2017;389: 1299-311.
7. Logan RF, Patnick J, Nickerson C, et al. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut 2012;61:1439-46.
8. Hurtado JL, Bacigalupe A, Calvo M, et al. Social inequalities in a population based colorectal cancer screening programme in the Basque Country. BMC Public Health 2015;15:1021.
9. 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;359:1291-300.

4

THE SECOND ROUND OF THE DUTCH COLORECTAL CANCER SCREENING PROGRAMME: IMPACT OF AN INCREASED FIT CUT-OFF LEVEL

Arthur I. Kooyker, **Esther Toes-Zoutendijk**, Annemieke W.J. Opstal-van Winden, Manon C.W. Spaander, Maaike Buskermolen, Anneke J. van Vuuren, Ernst J. Kuipers, Folkert J. van Kemenade, Christian Ramakers, Maarten G.J. Thomeer, Evelien Dekker, Iris D. Nagtegaal, Harry J. de Koning, Monique E. van Leerdam, Iris Lansdorp-Vogelaar

Submitted

ABSTRACT

Background

The Dutch colorectal cancer (CRC) screening programme started in 2014, inviting the target-population biennially to perform a faecal immunochemical test (FIT). We aimed to present the results of the second round (2016) and evaluate the impact of increasing the FIT cut-off halfway the first round (2014) on outcomes in the second round.

Methods

Prospectively collected data were obtained from the national screening information system (ScreenIT). Participants were categorised based on first-round participation status and used FIT cut-off level: previously tested with 15 µg (FIT^{1st-15}) or 47 µg (FIT^{1st-47}) and previous non-participants (FIT^{1st-np}). Second-round screening was done with a 47 µg Hb/g faeces FIT cut-off. Among others, outcomes included second-round participation and detection rate of advanced neoplasia (AN) and cumulative detection rate of AN over two rounds.

Results

348,891 (75.9%) out of 459,740 invitees participated in the second round. Participation rates were 93.4% among previous participants and 21.0% among previous non-participants. FIT^{1st-47} participants had a significantly higher detection rate of AN (15.3 vs. 10.4 per 1,000 participants) compared to FIT^{1st-15} participants in the second round, while their cumulative detection rate of AN over two rounds was significantly lower (45.6 vs. 52.6 per 1,000 participants).

Conclusions

Participation in the Dutch CRC screening programme was consistently high. Second-round detection rates depended on the first-round FIT cut-off. The cumulative detection over two rounds was higher among FIT^{1st-15} participants. These findings suggest that a substantial part, but not all, of the missed findings in the first round due to the increased FIT cut-off were detected in the subsequent round.

INTRODUCTION

Many countries have recently introduced colorectal cancer (CRC) screening with the aim to reduce CRC incidence and mortality. These programmes use different screening strategies.[1] Colonoscopy is the gold standard for detecting advanced neoplasia (AN) because of its high sensitivity. However, colonoscopy is an invasive procedure that demands extensive resources when used for primary screening on a population level. Many countries therefore prefer a non-invasive faecal test for primary screening, followed by colonoscopy when tested positive. Of the currently available faecal tests, faecal immunochemical testing (FIT) is associated with the highest participation and a high diagnostic performance.[2-6] Besides its superior characteristics compared to other faecal testing screen modalities, FIT offers the advantage to adjust the cut-off level to match local resources.[7] It allows for optimising the balance between the number of true- and false-positives, potentially impacting the detection rate.[8,9] Modelling studies based on real-life data reported that annual FIT at a low cut-off is equally effective in reducing CRC-related mortality as 10-yearly primary colonoscopy screening.[10] However, many organised programmes are currently forced to use a higher FIT cut-off and a longer screening interval, often due to a limited colonoscopy capacity.[1,11-13]

The Dutch FIT-based CRC screening programme started in 2014 after extensive piloting in previous years. During the first months after the start we observed a higher participation rate, a higher FIT positivity rate and a lower positive predictive value (PPV) compared to the results of the preceding pilot studies. Because the referral rate also exceeded the colonoscopy capacity, the FIT cut-off was increased halfway during the first year, resulting in a lower positivity rate and a higher PPV for advanced neoplasia.[14] In this study we evaluated participation in the second round and estimated the impact of the adjusted FIT cut-off in the first round on screening outcomes in the second round.

METHODS

The design of the Dutch CRC screening programme and its real-time monitoring system have previously been described.[14] In summary, the target population consists of individuals aged 55 to 75 years old, who are invited every two years to perform a FIT (FOB-Gold, Sentinel). Participants with a positive FIT (µg Hb/g faeces above the cut-off) are referred for a pre-colonoscopy intake. Eligible individuals are then scheduled for colonoscopy. The target population was invited gradually by birth-cohort.

Study population

In 2014, at the start of the screening programme, individuals that reached the age of 63, 65, 67, 75 and 76 years were invited for first-round screening. Persons aged 76 years were also

invited in 2014, due to a delayed implementation of the programme. For the second round, the same target group was re-invited in 2016, except for the invitees who tested positive in the first round, who had become older than 75 years, or who had deregistered permanently from the screening programme. In the first half year of 2014, a FIT cut-off of 15 µg Hb/g faeces was used. This was increased to 47 µg Hb/g faeces in the second half of 2014. In the second round, all FIT samples were analysed with FIT a cut-off of 47 µg Hb/g faeces.

Data collection

Of all invitees of the first round, data on participation status, FIT-result (µg Hb/g faeces), pre-colonoscopy intake and colonoscopy results in the first and/or second screening round were collected from the national screening information-system (ScreenIT).

Outcomes

Primary outcomes of this study were participation rate, FIT positivity rate, PPV for AN and detection rate of AN in the second screening round. An invitee was considered a participant when a FIT stool-sample was returned and a non-participant when there was no response or when the invitee deregistered. The participation rate was defined as the number of participants divided by the number of individuals invited. The positivity rate was defined as the number of participants with a FIT-result at or above the cut-off divided by the number of participants with an assessable FIT. The participation rate for pre-colonoscopy intake was defined as the number of persons who attended the pre-colonoscopy intake divided by the number of FIT-positives. The participation rate for colonoscopy was defined as the number of persons that underwent colonoscopy divided by the number of persons with a positive FIT. AN was considered a relevant finding within the CRC screening programme and was defined as CRC or advanced adenoma (AA).[15] AA was defined as any adenoma with histology showing 25% or greater villous component or high-grade dysplasia or an adenoma with size of 10 mm or larger. The PPV for AN was calculated as the number of persons detected with AN divided by the number of persons who underwent a colonoscopy. The detection rate of AN was defined as the number of persons detected with AN of those who returned an assessable FIT. Secondary outcomes were cumulative positivity rate, cumulative detection rate of AN, number needed to scope (NNScope) to detect AN in one patient, and the association between concentration µg Hb/g faeces (FIT cut-off level) in the first round and screening-outcomes in the second round. Cumulative positivity rate was defined as the number of positive FIT results over both rounds divided by the number of participants that returned an assessable FIT in both rounds or tested positive in the first round. The cumulative detection rate was defined as the number of CRC or AN detected over both rounds in participants that returned an assessable FIT in both rounds or tested positive in the first round. The NNScope was defined as the number of performed colonoscopies divided by the number of detected CRC or AN, over both rounds. The FIT-level in the first round was tested on the association with positive FIT-result, PPV and detection of AN in the second round.

Analyses

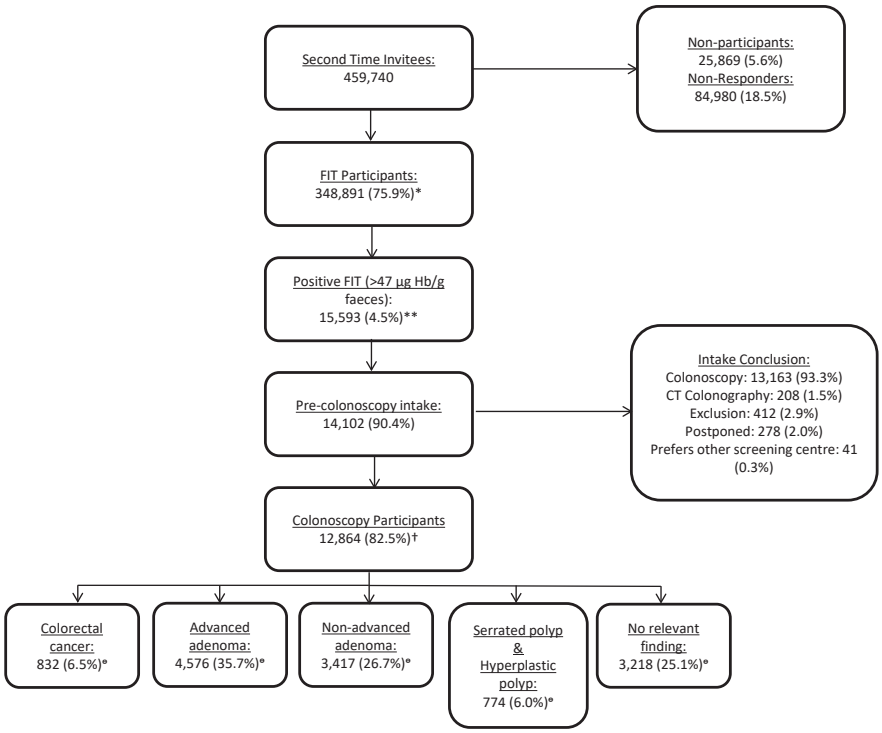
All invitees to the second round were analysed for participation rate, positivity rate, participation rate of pre-colonoscopy intake and colonoscopy, PPV for AN and detection rate of AN. Primary outcomes were presented for three different subgroups: individuals that were tested with a FIT cut-off of 15 µg Hb/g faeces in the first round ($FIT^{1st\text{-}15}$), individuals that were tested with a FIT cut-off of 47 µg Hb/g faeces in the first round ($FIT^{1st\text{-}47}$) and individuals that did not participate in the first round ($FIT^{1st\text{-}np}$). Secondary outcomes excluded participants of the first round of the oldest birth cohorts (1938 and 1939), since those participants were not part of the target population of 2016 and therefore not re-invited. The cumulative positivity rate, cumulative detection rate and NNScope over two screen-rounds were compared between $FIT^{1st\text{-}15}$ and $FIT^{1st\text{-}47}$ participants of both rounds. To assess the association between FIT-level in the first round and screening outcomes in the second round, the FIT-results of first round participants were categorised in 0 µg, 1-14 µg and 15-46 µg Hb/g faeces (the latter group was only applicable to $FIT^{1st\text{-}47}$ participants).

Descriptive statistics were computed of the primary outcomes, including 95% confidence intervals (95%CI). Differences between groups were tested for statistical significance ($p < 0.05$) using the Chi-square test or the Student's t-test. Age-adjusted rates for the primary outcomes and secondary outcomes were calculated for groups with more than 50 invitees, therefore only age groups of 65, 67 and 69 years old were included. Because of substantially different age-distributions between subgroups, the chi-square was not applicable. Instead, differences between the age-adjusted rates were tested with the Standardised Rate Ratio (SRR).[16,17] If the 95%CI of the SRR includes 1, no significant difference was observed between the age-adjusted rates. Multivariable logistic regression was performed to estimate the odds ratios (ORs) of the FIT-level in the first round on screening outcomes in the second round, adjusted for gender and age. As sensitivity analyses, we 1) tested the outcomes on significant differences between gender and 2) only included $FIT^{1st\text{-}47}$ participants to rule out verification bias, since only FIT-positive participants are referred for colonoscopy (Appendix I & II).

RESULTS

In total, 459,740 individuals were invited for second-round screening. Of those, 348,891 (75.9%) participated and 4.5% tested FIT-positive (Figure 1 & Table 1). Out of the 14,102 (90.4%, 95%CI: 90.0-90.9%) individuals that attended the pre-colonoscopy intake, 93.3% (95%CI: 92.9-93.7%) were advised to undergo colonoscopy. Of all individuals that tested FIT-positive, 12,864 (82.5%, 95%CI: 81.9-83.1%) participants underwent colonoscopy at which 832 CRCs and 4,576 AAs were detected. This resulted in a detection rate of 15.5 (95%CI: 15.1-15.9) AN per 1,000 participants.

Figure 1: Flow chart of the second round of the Dutch CRC screening programme



Abbreviations: FIT (faecal immunochemical testing)
* Off all participants, 348,726 (99.95%) individuals returned an assessable FIT
** Divided by the number of assessable FITs
†Divided by the number of participants with a positive FIT
ϴ Divided by the number of participants with a confirmed conclusion of colonoscopy (n = 12,817)

Participation

Of all 348,071 second-round invitees who participated in the first round, 325,392 (93.5%) also participated in the second round (Table 2). Of all 111,669 second-round invitees who did not participate in the first round, a total of 23,499 (21.0%) participated in the second round.

Yield of screening

Among second-round invitees that had also participated in the first round, there were 39,257 $FIT^{1st\text{-}15}$ and 286,135 $FIT^{1st\text{-}47}$ participants (Table 2). $FIT^{1st\text{-}15}$ participants of the second round had a positivity rate of 3.3%, a PPV for AN of 36.9% (95%CI: 34.1-39.8%), and a detection rate of AN of 10.4 (95%CI: 9.4-11.4) per 1,000 participants. $FIT^{1st\text{-}47}$ participants had in the second round a higher positivity rate of 4.3%, a higher PPV for AN of 41.2% (95%CI: 40.2-42.1%), and a higher detection rate of AN of 15.0 (95%CI: 14.6-15.5) per 1,000 participants. SRR presented a significantly higher age-adjusted positivity rate (SRR: 1.3, 95%CI: 1.1-1.5) and

Table 1: Total yield of the second round

	Participation rate		Positivity rate		PPV CRC		PPV AA		Detection rate CRC		Detection rate AA	
	n	%	n	%	n	%	n	%	n	‰	n	‰
Total	348,891	75.9 (75.8 - 76.0)	15,593	4.5 (4.4 - 4.5)	832	6.5 (6.1 - 6.9)	4,576	35.7 (34.9 - 36.5)	832	2.4 (2.2 - 2.6)	4,576	13.1 (12.7 - 13.5)
Gender												
Men	168,916	74.0 (73.8 - 74.2)	8,998	5.3 (5.2 - 5.4)	495	6.7 (6.1 - 7.2)	2,878	38.7 (37.6 - 39.8)	495	2.9 (2.7 - 3.2)	2,878	17.0 (16.4 - 17.7)
Women	179,975	77.8 (77.6 - 77.9)	6,595	3.7 (3.6 - 3.8)	337	6.3 (5.6 - 6.9)	1,698	31.6 (30.3 - 32.8)	337	1.9 (1.7 - 2.1)	1,698	9.4 (9.0 - 9.9)
Age												
62	26,027	75.9 (72.4 - 73.3)	1,064	4.1 (3.9 - 4.3)	40	4.6 (3.4 - 6.2)	296	34.3 (31.2 - 37.5)	40	1.5 (1.1 - 2.1)	296	11.4 (10.2 - 12.7)
65	84,442	76.3 (76.0 - 76.5)	3,572	4.2 (4.1 - 4.4)	186	6.3 (5.5 - 7.2)	1,048	35.4 (33.7 - 37.1)	186	2.2 (1.9 - 2.5)	1,048	12.4 (11.7 - 13.2)
67	108,248	76.1 (75.8 - 76.3)	4,679	4.3 (4.2 - 4.4)	222	5.8 (5.1 - 6.5)	1,360	35.3 (33.8 - 36.8)	222	2.1 (1.8 - 2.3)	1,360	12.6 (11.9 - 13.3)
69	130,174	76.1 (75.9 - 76.3)	6,278	4.8 (4.7 - 4.9)	384	7.5 (6.8 - 8.2)	1,872	36.5 (35.2 - 37.8)	384	3.0 (2.7 - 3.3)	1,872	14.4 (13.8 - 15.0)

Abbreviations: PPV (Positive Predictive Value); AA (Advanced Adenoma)

Participants older than 75 years old (n=14) are excluded from analyses. Because of small numbers, the age categories of 66 (n = 1) and 72 years old (n = 1) is not shown.

Table 2: Outcomes of the second screening round in first round participants (tested with a cut-off of 15 µg Hb/g faeces or 47 µg Hb/g faeces) and in first round non-participants

	Participation rate		Positivity rate		PPV CRC		PPV AA		Detection rate CRC		Detection rate AA	
	Previous participants											
	n	%	n	%	n	%	n	%	n	‰	n	‰
Total	325,392	93.5 (93.4 - 93.6)	13,743	4.2 (4.2 - 4.3)	708	6.1 (5.7 - 6.6)	3,996	34.6 (33.8 - 35.5)	708	2.2 (2.0 - 2.3)	3,996	12.3 (11.9 - 12.7)
Adjusted¶		*93.5*		*4.2*		*6.1*		*34.6*		*2.2*		*12.3*
FIT 1st-15												
Total	39,257	94.1 (93.8 - 94.3)	1,303	3.3 (3.1 - 3.5)	63	5.7 (4.5 - 7.2)	344	31.2 (28.5 - 34.0)	63	1.6 (1.3 - 2.1)	344	8.8 (7.9 - 9.7)
Adjusted¶				*3.4**		*5.5*		*30.5*		*1.6*		*8.8**
Gender												
Men	18,497	94.0 (93.7 - 94.4)	732	4.0 (3.7 - 4.2)	36	5.8 (4.2 - 7.9)	209	33.7 (30.0 - 37.5)	36	1.9 (1.4 - 2.7)	209	11.3 (9.9 - 12.9)
Women	20,760	94.0 (93.7 - 94.3)	571	2.8 (2.5 - 3.0)	27	5.6 (3.9 - 8.0)	135	28.0 (24.2 - 32.2)	27	1.3 (0.9 - 1.9)	135	6.5 (5.5 - 7.7)
Age												
65	1,953	94.5 (93.5 - 95.4)	67	3.4 (2.7 - 4.3)	2	3.6 (1.0 - 12.1)	15	26.8 (17.0 - 39.6)	2	1.0 (0.3 - 3.7)	15	7.7 (4.7 - 12.6)
67	23,874	94.0 (93.7 - 94.3)	741	3.1 (2.9 - 3.3)	32	5.1 (3.7 - 7.2)	191	30.7 (27.2 - 34.4)	32	1.3 (0.9 - 1.9)	191	8.0 (6.9 - 9.2)
69	13,430	94.1 (93.7 - 94.5)	495	3.7 (3.4 - 4.0)	29	6.8 (4.8 - 9.6)	138	32.5 (28.2 - 37.1)	29	2.2 (1.5 - 3.1)	138	10.3 (8.7 - 12.1)
FIT 1st-47												
Total	286,135	93.4 (93.3 - 93.5)	12,440	4.3 (4.3 - 4.4)	645	6.2 (5.7 - 6.7)	3,652	35.0 (34.1 - 35.9)	645	2.3 (2.1 - 2.4)	3,652	12.8 (12.4 - 13.2)
Adjusted¶				*4.4**		*6.3*		*35.2*		*2.3*		*13.0**
Gender												
Men	138,117	93.4 (93.3 - 93.5)	7,108	5.1 (5.0 - 5.3)	373	6.2 (5.6 - 6.9)	2,269	37.9 (36.7 - 39.1)	373	2.7 (2.4 - 3.0)	2,269	16.4 (15.8 - 17.1)
Women	148,018	93.4 (93.2 - 93.5)	5,332	3.6 (3.5 - 3.7)	272	6.1 (5.5 - 6.9)	1,383	31.1 (29.8 - 32.5)	272	1.8 (1.6 - 2.1)	1,383	9.3 (8.9 - 9.8)
Age												

Table 2: Outcomes of the second screening round in first round participants (tested with a cut-off of 15 µg Hb/g faeces or 47 µg Hb/g faeces) and in first round non-participants (continued)

	Participation rate		Positivity rate		PPV CRC		PPV AA		Detection rate CRC		Detection rate AA	
					Previous participants							
	n	%	n	%	n	%	n	%	n	‰	n	‰
62	23,875	93.0 (92.7 - 93.3)	910	3.8 (3.6 - 4.1)	36	4.8 (3.5 - 6.5)	248	32.9 (29.7 - 36.4)	36	1.5 (1.1 - 2.1)	248	10.4 (9.2 - 11.8)
65	76,776	93.6 (93.4 - 93.8)	3,087	4.0 (3.9 - 4.2)	152	5.8 (5.0 - 6.8)	895	34.3 (32.5 - 36.2)	152	2.0 (1.7 - 2.3)	895	11.7 (10.9 - 12.4)
67	77,209	93.5 (93.3 - 93.6)	3,387	4.4 (4.2 - 4.5)	151	5.3 (4.5 - 6.2)	994	34.8 (33.1 - 36.6)	151	2.0 (1.7 - 2.3)	994	12.9 (12.1 - 13.7)
69	108,275	93.3 (93.2 - 93.5)	5,056	4.7 (4.5 - 7.8)	306	7.3 (6.5 - 8.1)	1,515	35.9 (34.5 - 37.4)	306	2.8 (2.5 - 3.2)	1,515	14.0 (13.3 - 14.7)
Previous non-participants												
-	n	%	n	%	n	%	n	%	n	‰	n	‰
Total	23,499	21.0 (20.8 - 21.3)	1,850	7.9 (7.6 - 8.3)	124	9.7 (8.2 - 11.4)	580	45.2 (42.5 - 48.0)	124	5.3 (4.5 - 6.3)	580	24.8 (22.8 - 26.8)
Adjusted¶		*21.0***		*7.9***		*9.8***		*45.3***		*5.4***		*24.8***
Gender												
Men	12,302	20.3 (19.9 - 20.6)	1,158	9.4 (8.9 - 10.0)	86	10.4 (8.5 - 12.7)	400	48.5 (45.1 - 52.0)	86	7.0 (5.7 - 8.7)	400	32.6 (29.6 - 35.9)
Women	11,197	22.0 (21.6 - 22.3)	692	6.2 (5.8 - 6.7)	38	8.3 (6.1 - 11.2)	180	39.3 (34.9 - 43.8)	38	3.4 (2.5 - 4.7)	180	16.1 (14.0 - 18.6)
Age												
62	2,152	21.4 (20.6 - 22.2)	154	7.2 (6.2 - 8.3)	4	3.6 (1.4 - 9.0)	48	43.6 (34.7 - 53.0)	4	1.9 (0.7 - 4.8)	48	22.4 (16.9 - 29.5)
65	5,713	21.5 (21.0 - 22.0)	418	7.3 (6.7 - 8.1)	32	10.7 (7.7 - 14.8)	138	46.3 (40.7 - 52.0)	32	5.6 (4.0 - 7.9)	138	24.2 (20.6 - 28.6)
67	7,165	20.9 (20.5 - 21.3)	551	7.7 (7.1 - 8.4)	39	10.2 (7.5 - 13.6)	175	45.7 (40.8 - 50.7)	39	5.5 (4.0 - 7.5)	175	24.5 (21.2 - 28.4)
69	8,469	20.8 (20.4 - 21.2)	727	8.6 (8.0 - 9.2)	49	10.0 (7.6 - 12.9)	219	44.6 (40.3 - 49.0)	49	5.8 (4.4 - 7.7)	219	25.9 (22.7 - 29.5)

Abbreviations: PPV (Positive Predictive Value); AA (Advanced Adenoma)

¶ age-adjusted rates; *Significant (α = 0.05) difference between used FIT cut-offs in the first round; **Significantly (α = 0.05) different compared to previous participants

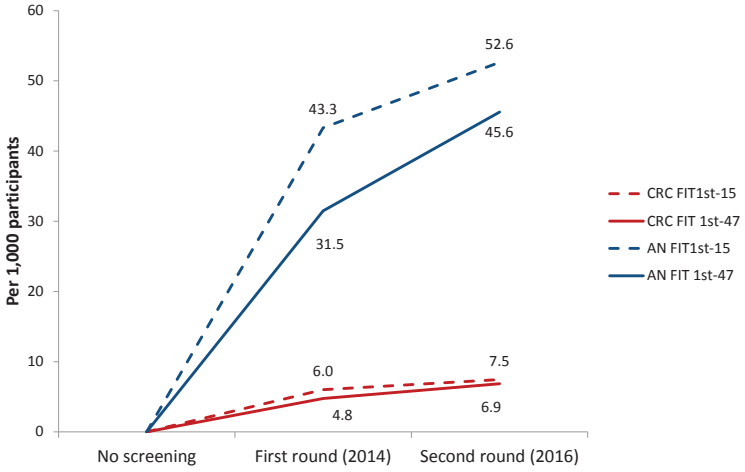
Because of small numbers, the results of the age categories of 66 (n = 1) and 72 (n = 1) years old in the non-participants of the first round are not shown.

age-adjusted detection rate of AN (SRR: 1.5, 95%CI: 1.2-1.9) in FIT^{1st-47} participants of the second round (Table 2). Differences in age-adjusted PPV between the used cut-offs were non-significant. The FIT^{1st-np} participants of the second round showed a high positivity rate of 7.9%, PPV of 54.9% (95%CI: 52.2-57.6%) for AN and detection rate of AN of 30.1 (95%CI: 27.9-32.3) per 1,000 participants (Table 2). All outcomes were significantly higher compared to second-round participants that had participated in the first round.

Cumulative rates and NNScope

Of FIT^{1st-15} participants, 13.1% (95%CI: 12.8-13.4%) tested positive in the first (10.1%) or second (3.0%) round (cumulative positivity rate). This cumulative positivity rate was higher compared to the FIT^{1st-47} participants, of which 10.4% (95%CI: 10.3-10.5%) tested positive in the first (6.3%) or second (4.1%) round. The age-adjusted cumulative positivity rate over two rounds was significantly higher in FIT^{1st-15} participants (SRR: 1.2, 95%CI: 1.2-1.3). Per 1,000 FIT^{1st-15} participants, CRC was detected in 7.5 (95%CI: 6.7-8.3) participants in the first (6.0) or second (1.4) round and AN were detected in 52.6 (95%CI: 50.6-54.8) participants in the first (43.3) or second (9.3) round (cumulative detection rate) (Figure 2). A lower cumulative detection rate of CRC and AN was observed per 1,000 FIT^{1st-47} participants, in which 6.9 (95%CI: 6.6-7.2) participants were detected with CRC in the first (4.8) or second (2.1) round and 45.6 (95%CI: 44.8-46.3) with AN in the first (31.5) or second (14.1) round. Age-adjusted rates showed a non-significant difference in the cumulative detection rate over two rounds for CRC (SRR: 1.1, 95%CI: 0.8-1.4), yet significantly more FIT^{1st-15} participants were detected with AN (SRR: 1.2, 95%CI: 1.1-1.3). Over two rounds, 14.9 (95%CI: 14.5-15.3) FIT^{1st-15} participants and 12.7 (95%CI:

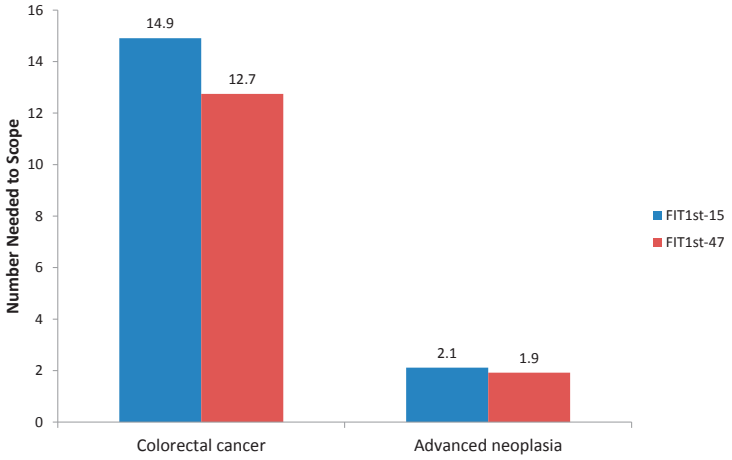
Figure 2: Cumulative detection rate of colorectal cancer and advanced neoplasia over two rounds of screening



Abbreviations: CRC (colorectal cancer), AN (advanced neoplasia)

12.6 – 12.9) FIT^{1st-47} participants needed to undergo colonoscopy (NNScope) to detect CRC in one participant, with a significant difference after age-adjustment (SRR: 1.2, 95%CI: 1.1-1.3). A similar pattern was observed for the NNScope to detect AN in one participant (Figure 3).

Figure 3: Number of participants that underwent colonoscopy (NNScope) over two rounds of screening to detect one colorectal cancer or advanced neoplasia



Yield by previous FIT result

The positivity rate, PPV and detection rate in the second round was strongly related to the concentration of Hb detected in faeces in the first round. Participants with a FIT-level in the first round between 15 and 47 µg Hb/g faeces showed a positivity rate of 23.3%, a PPV for AN of 60.3% and a detection rate of AN of 120.3 per 1,000 participants in the second round (Table 3). The outcomes in these participants were significantly higher than in participants with a FIT-level below 15 µg in the first round. Compared to participants with no (0 µg) detectable Hb in their faeces sample, a participant with a FIT-level between 15 and 47 µg was remarkably more likely to test FIT-positive (OR 11.9, 95%CI: 11.3-12.5) or have an AN detected during colonoscopy (OR 23.2, 95%CI: 21.5-25.1). While the sensitivity analyses ruled out verification bias, it pointed out significant differences between male and female participants (Appendix I & II). However, this did not change the conclusion, as both genders presented a strong correlation between first round FIT-result and outcomes in the subsequent round.

Table 3: Yield of the second round relative to FIT results of the first round

First screening round FIT result (Hb/g faeces)	0 µg	>0 µg and <15 µg	≥15 µg and <47 µg	*p* value
Total	248,310	66,030	10,961	
Positivity rate				
n (%)	6,000 (2.4)	5,187 (7.9)	2,553 (23.3)	< 0.001
OR (95% CI)	-	3.4 (3.3 - 3.5)*	11.9 (11.3 - 12.5)*	
PPV AN				
n (%)	1,390 (27.9)	1,993 (45.7)	1,319 (60.3)	< 0.001
OR (95% CI)	-	2.2 (2.0 - 2.4)	3.9 (3.5 - 4.3)	
Detection rate AN				
n (per 1,000 participants)	1,390 (5.6)	1,993 (30.2)	1,319 (120.3)	< 0.001
OR (95% CI)	-	5.4 (5.0 - 5.8)*	23.2 (21.5 - 25.1)*	

Abbreviations: OR (Odds Ratio), PPV (Positive Predictive Value); AN (Advanced Neoplasia)
* Significant interaction between male and female gender (see appendix)
OR are adjusted for age and gender

DISCUSSION

This study presents the results of the second round of the Dutch FIT-based CRC screening programme and evaluates the impact of increasing the FIT cut-off in the first round on the yield of the second round. We observed a consistently high participation as almost all first-round participants also participated in the second round. The detection rate of AN in the second round was significantly higher in FIT^{1st-47} participants compared to FIT^{1st-15} participants. Nevertheless, the cumulative detection rate of AN over two rounds was significantly lower in FIT^{1st-47} participants. We found a strong correlation between the concentration µg Hb/g faeces in the first round and the detection of AN in the subsequent round.

The Dutch government was the first to change the FIT cut-off in a running national CRC screening programme. Therefore we are the only country yet in which evaluation of the programme can demonstrate the impact of adjusting the FIT cut-off on the yield at a population level. The main reason to increase the FIT cut-off during the first round was to reduce colonoscopy demand and the proportion of false-positive FIT-results. As previously reported, this indeed successfully decreased the positivity rate and increased the PPV in first round FIT^{1st-47} participants, at the cost of a lower detection rate (14). We demonstrated in the current study that, cumulatively, the detection rate over two rounds was still lower in FIT^{1st-47} participants. However, we also observed that the difference in cumulative detection rate of AN between FIT^{1st-47} and FIT^{1st-15} participants decreased from 11.8 AN per 1,000 participants in the first round, to 7.0 AN per 1,000 participants in the second round. This means that a substantial part of the missed lesions in the first screening-round due to the increased FIT cut-off from 15 to 47 µg Hb/g faeces are detected at the subsequent round. If the difference

in cumulative detection rate of AN between FIT^{1st-47} and FIT^{1st-15} participants keeps declining per subsequent screening, the impact of adjusting the FIT cut-off on yield of screening might become insignificant within one or two subsequent screening rounds.

The majority of undetected CRCs due to the increased FIT cut-off seem to be detected in the second round, as the cumulative detection rate of CRC over two rounds in FIT^{1st-47} and FIT^{1st-15} participants was almost similar. This suggests that increasing the FIT cut-off resulted in limited harms. However, it is unclear if the higher cut-off resulted in more interval cancers and if the detected CRCs in the second round were still diagnosed in an early stage, which is a prerequisite for screening to be effective. Therefore, interval cancers have to be analysed and data on stage distribution have to be evaluated. The effect of increasing the cut-off in the older birth cohorts, 1938 and 1939, which were not re-invited for a second round as they exceeded the target age, should be assessed separately. As the missed lesions in FIT^{1st-47} participants of these birth cohorts will not be detected in a subsequent round, they might potentially progress to symptomatic CRC.

The reported participation rate (75.9%) in the Netherlands can be considered the highest in a subsequent round, even higher than observed in the Dutch CRC screening pilots.[18-20] However, this outcome is overestimated by approximately 2%, because the calculation of participation rate excluded invitees who deregistered permanently during the first round and were therefore not invited to the second round.

The strong correlation between the level of Hb concentration of a negative FIT result and the chance of detecting AN in the subsequent round, suggests an excellent opportunity for more personalised FIT screening based on Hb concentration. For example, the screening interval of participants with no detectable blood (0 µg Hb/g faeces) could potentially be prolonged. Similarly, it may be reasonable to re-invite participants that tested just negative (15 – 47 µg Hb/g faeces) already after one year to detect relevant findings in time. In concordance with our data, a similar association between FIT level and outcomes in the subsequent rounds has previously been presented by a Dutch, Spanish and Taiwanese study.[21-23]

Important strengths of our study are the nation-wide implementation of the screening programme, the large sample size and the well-developed registration, therefore providing accurate data. Nevertheless, two limitations are noteworthy. As mentioned before, data on stage distribution are lacking, which would provide important information on the potential consequences of missed lesions in the first round caused by the increased FIT cut-off. These data will be available in the near future. Second, not every birth cohort has been invited yet for the Dutch CRC screening programme; conclusions are therefore based on a few age groups.

Notwithstanding these limitations, we are the first to present on how using a different FIT cut-off in the first round impacts the outcomes of a subsequent round. Our findings are of value to other FIT-based CRC screening programmes considering an appropriate cut-off in their setting, in particular when the used FIT cut-off is within the same range.

In conclusion, participation in the Dutch CRC screening programme was high and consistent. Our results showed that using a higher FIT cut-off in CRC screening has limited impact on the yield of screening because a substantial part of AN will be detected in subsequent rounds. To confirm whether these AN are still detected in an early stage, retrieving more information on the stage distribution of CRCs detected in the second round is important.

REFERENCES

1. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-1649.
2. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus Faecal Immunochemical Testing in Colorectal-Cancer Screening. N Engl J Med. 2012;3668):697-706.
3. Hol L, Wilschut J, 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;100:1103-1110.
4. 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;58:241-248.
5. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59(1):62-68.
6. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random Comparison of Guaiac and Immunochemical Faecal Occult Blood Tests for Colorectal Cancer in a Screening Population. Gastroenterology. 2008;135(1):82-90.
7. Wilschut JA, Habbema JDF, Van Leerdam ME, et al. Faecal occult blood testing when colonoscopy capacity is limited. J Natl Cancer Inst. 2011;103(23):1741-1751.
8. De Wijkerslooth TR, Stoop EM, Bossuyt PM, et al. Immunochemical faecal occult blood testing is equally sensitive for proximal and distal advanced Neoplasia. Am J Gastroenterol. 2012;107(10):1570-1578.
9. Haug U, Hundt S, Brenner H. Quantitative immunochemical faecal occult blood testing for colorectal adenoma detection: Evaluation in the target population of screening and comparison with qualitative tests. Am J Gastroenterol. 2010;105(3):682-690.
10. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of benefits, burden, and harms of colorectal cancer screening strategies: Modeling study for the US preventive services Task Force. JAMA. 2016;315(23):2595-2609.
11. Grazzini G, Visioli CB, Zorzi M, et al. Immunochemical faecal occult blood test: Number of samples and positivity cutoff. What is the best strategy for colorectal cancer screening? Br J Cancer. 2009;100(2):259-265.
12. UK National Screening Committee Screening. UK National Screening Committee Screening for Bowel Cancer. 2015.
13. Aniwan S, Ratanachu EKT, Pongprasobchai S, et al. The Optimal Cut-Off Level of The Faecal Immunochemical Test For Colorectal Cancer Screening in a Country with Limited Colonoscopy Resources: A Multi-Center Study from Thailand. Asian Pac J Cancer Prev. 2017;18(2):405-412.
14. Toes-Zoutendijk E, van Leerdam ME, Dekker E, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Programmeme and Optimization by Altering Faecal Immunochemical Test Cut-Off Levels. Gastroenterology. 2017;152(4):767-775.e2.
15. 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(SUPPL3).
16. Breslow N, Day N. Statistical Methods in Cancer Research. New York: Oxford University Press; 1987.
17. Boyle P, Parkin D. Statistical Methods for Registries, In: Cancer Registration: Principles and Methods; International Agency for Research on Cancer, Lyon, p. 126-158; 1991.
18. Van Der Vlugt M, Grobbee EJ, Bossuyt PMM, et al. Adherence to colorectal cancer screening: Four rounds of faecal

immunochemical test-based screening. Br J Cancer. 2017;116(1):44-49.

19. Kapidzic A, Grobbee EJ, Hol L, et al. Attendance and yield over three rounds of population-based faecal immunochemical test screening. Am J Gastroenterol. 2014;109(8):1257-1264.
20. Stegeman I, Van Doorn SC, Mundt MW, et al. Participation, yield, and interval carcinomas in three rounds of biennial FIT-based colorectal cancer screening. Cancer Epidemiol. 2015;39:388-393.
21. Grobbee EJ, Schreuders EH, Hansen BE, et al. Association Between Concentrations of Hemoglobin Determined by Faecal Immunochemical Tests and Long-term Development of Advanced Colorectal Neoplasia. Gastroenterology. 2017;153(5):1251-1259.e2.
22. Buron A, Román M, Augé JM, et al. Changes in FIT values below the threshold of positivity and short-term risk of advanced colorectal neoplasia: Results from a population-based cancer screening programme Eur J Cancer. 2019;107:53-59.
23. Chen LS, Yen AMF, Chiu SYH, 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):551-558.

APPENDIX I

Table 1: FIT positivity rate and detection rate for advanced neoplasia in the second screening round relative to the first screening round FIT result by sex

First screening round FIT result (Hb/g faeces)	0 µg	>0 µg and <15 µg	≥15 µg and <47 µg
Positivity rate (OR (95% CI))			
Male	-	3.2 (3.0 - 3.4)	11.3 (10.5 - 12.0)
Female	-	3.6 (3.4 - 3.8)	12.8 (11.8 - 13.9)
Detection rate AN (OR (95% CI))			
Male	-	5.0 (4.6 - 5.4)	20.6 (18.6 - 22.7)
Female	-	6.0 (5.4 - 6.7)	28.2 (24.8 - 31.9)

Note: other outcomes were not significantly different between male and female
Abbreviations: OR (Odds Ratio); AN (Advanced neoplasia)

APPENDIX II

Table 1: Yield of the second round relative to FIT results of the first round, including only participants tested in the first round with 47 µg Hb/g faeces

First screening round FIT result (Hb/g faeces)	0 µg	>0 µg and <15 µg	≥15 µg and <47 µg	*p* value
Total	217,435	57,662	10,948	
Positivity rate				
n (%)	5,334 (2.5)	4,553 (7.9)	2,550 (23.3)	< 0.001
Odds-ratio (95% CI)	-	3.4 (3.2 - 3.5)*	11.7 (11.1 - 12.3)*	
PPV AN				
n (%)	1,230 (27.8)	1,748 (45.7)	1,317 (60.3)	< 0.001
Odds-ratio (95% CI)	-	2.2 (2.0 - 2.4)	3.9 (3.5 - 4.3)	
Detection rate AN				
n (per 1,000 participants)	1,230 (5.7)	1,748 (30.3)	1,317 (120.3)	< 0.001
Odds-ratio (95% CI)	-	5.4 (5.0 - 5.8)*	22.9 (21.2 - 24.9)*	

Outcomes are adjusted for age and sex
Abbreviations: PPV (Positive Predictive Value); AN (Advanced Neoplasia)

5

INCIDENCE OF INTERVAL COLORECTAL CANCER AFTER NEGATIVE RESULT FROM FIRST-ROUND FAECAL IMMUNOCHEMICAL SCREENING TESTS, BY CUTOFF VALUE AND PARTICIPANT SEX AND AGE

Esther Toes-Zoutendijk, Arthur I. Kooyker, Evelien Dekker, Manon C.W. Spaander, Annemieke W.J. Opstal-van Winden, Christian Ramakers, Maaike Buskermolen, Anneke J. van Vuuren, Ernst J. Kuipers, Folkert J. van Kemenade, Marie-Louise F. Velthuysen, Maarten G.J. Thomeer, Harriët van Veldhuizen, Marjolein van Ballegooijen, Iris D. Nagtegaal, Harry J. de Koning, Monique E. van Leerdam*, Iris Lansdorp-Vogelaar*, on behalf of the Dutch national colorectal cancer screening working group.

* both authors contributed equally

Clinical Gastroenterology and Hepatology, 2019

ABSTRACT

Background

This study evaluated the interval cancer incidence after the first screening round in an organised colorectal cancer (CRC) screening programme using the FOB-Gold faecal immunochemical test (FIT) in relation to FIT cut-off.

Methods

Screening participants with a negative FIT in the first screening round in national population-based CRC screening programme in the Netherlands in 2014 were included in the study. Cumulative incidence of interval cancer after negative FIT and FIT sensitivity for CRC at a low (15 µg Hb/g faeces) and higher (47 µg Hb/g faeces) cut-off were estimated.

Results

Among the 485,112 participants with a negative FIT in 2014, 544 interval cancers were detected: 126 interval cancers among 111,800 FIT negatives at low cut-off and 418 interval cancers among 373,312 FIT negatives at higher cut-off. Mean age of individuals tested at the low cut-off was 72.0 years and at the higher cut-off was 66.7 years. The age-adjusted two-year cumulative incidences of interval cancer after negative FIT were 9.5 vs 13.8 per 10,000 persons at respectively the low cut-off and higher cut-off, which were statistically different (p 0.005). Age-adjusted FIT sensitivities for CRC were statistically different with 90.5% respectively 82.9% at the low and higher cut-off (p <0.0001). The overall FIT sensitivity for CRC was 87.4% among men and 82.6% among women (p <0.001).

Conclusions

The incidence of interval CRC after a negative FIT is low. Although FIT sensitivity for declined with a higher cut-off, it remained above 80%.

INTRODUCTION

Many countries have introduced a screening programme for colorectal cancer (CRC) in recent years. Different screening modalities are suitable for that purpose. Opportunistic screening programmes most often use colonoscopy for primary screening, while organised population-based programmes mostly prefer faecal immunochemical testing (FIT).[1] Colonoscopy has better test characteristics compared to FIT when applied for one-time screening, yet is invasive, burdensome, and costly. FIT is non-invasive, non-burdensome, and less costly, but has lower test sensitivity.[2-4] For optimal programme sensitivity and preventive effect, FIT should be repeated regularly.

FIT has been shown to be effective in detecting CRC at low cut-offs or short screening intervals.[5,6] Modelling studies suggested that by repeating FIT annually, with an assumed test sensitivity of 73.8% for CRC, the long-term preventive effect would be similar to colonoscopy screening.[5] The number of interval cancers in the Dutch CRC screening pilot study was recently evaluated, based on three biennial FIT screening rounds. This relatively small study showed an interval CRC incidence rate of 0.1% and a sensitivity of 77% over three screening rounds.[6] However, these interval CRCs were observed while using a very low FIT cut-off of 10 µg Hb/g faeces. In many regional or national population-based organised programmes a higher cut-off for a positive FIT with referral to colonoscopy is chosen for a better balance between true and false positives.[1] Six months after the start of the Dutch national programme, the FIT cut-off was increased from 15 to 47 µg Hb/g faeces, because of a higher than expected positivity rate with an associated lower positive predictive value and shortage in colonoscopy capacity.[7] Consequently, we assumed that 12% of the CRCs would be missed.[8]

Evaluation of the number of interval CRCs within organised population-based screening programmes is important. The results of the Dutch CRC FIT-based screening programme enable us to evaluate the number of interval CRCs after the first screening and determine the impact of using a relatively high versus a low FIT cut-off on the cumulative incidence and sensitivity of FIT for CRC.

METHODS

Screening programme and population

In the Netherlands a national population-based CRC screening programme was implemented in 2014, with biennial FIT screening for persons aged 55 through 75 years. The programme was rolled-out in 5 years (2014-2018), with a phased-implementation by age groups (birth cohorts). In 2014 individuals aged 60, 63, 65, 67, 75, and 76 years old were invited. For once also persons aged 76 years were invited in 2014, because the start of the programme was delayed. Individuals received an invitation letter by postal mail including one single FIT

(FOB-gold, Sentinel, Italy). Participants with a positive FIT were referred for colonoscopy. Participants with a negative FIT were re-invited 24 months after the previous invitation date. Note, this is not 24 months after a negative FIT, therefore screening interval could be shorter than 2 years. At the start in 2014, the cut-off for a positive test was defined at 15 µg Hb/g faeces. As a result of a higher than expected participation and positivity rate and a lower than expected positive predictive value for CRC and advanced adenomas (AA), it was decided to increase the cut-off in June 2014 to 47 µg Hb/g faeces. A more extensive description of the Dutch national CRC screening programme and the decision analysis on increasing the cut-off was given in a previous publication.[7] This current paper evaluated the interval CRCs of participants invited in the first year of the national Dutch CRC screening programme in 2014.

Outcomes

We estimated the cumulative incidence of interval cancers and test sensitivity. The cumulative incidence was calculated as the number of interval CRCs within two years after a negative FIT in the first screening round divided by the total number of individuals with a negative FIT in the first screening round. Number was presented per 10,000 individuals with a negative FIT. FIT sensitivity was approximated by the number of screen-detected CRCs after a positive FIT in the first screening round divided by the sum of screen-detected and interval CRCs in the first screening round. This is a commonly applied approximation in screening literature.

We defined FIT interval CRCs according to the internationally recommended nomenclature of the working group on interval CRC of the World Endoscopy Organization.[9] They designated an interval CRC as a CRC after a negative FIT but before the invitation of subsequent screening round with FIT.

An interval CRC in this study population was defined as follows for two distinct subgroups:

1) *Participants with a negative FIT in 2014 and eligible for screening in the subsequent round*: CRCs that occur between date of FIT analyses with negative FIT and date of invitation of the subsequent screening round.
2) *Participants with a negative FIT in 2014 and not eligible for screening in the subsequent round because of the upper age limit:* CRCs that occur between date of FIT analyses with negative FIT plus 24 months.

Screen-detected CRCs were defined as cancers detected within 6 months after a positive FIT in the first screening round.

Data collection

Data of participants with a negative FIT in 2014 were obtained from the national screening database ScreenIT; Hb concentration, gender, age, invitation date, and date of analyses. All individual records of these participants were sent to the Netherlands Cancer Registry (NCR). This registry contained information on cancers detected in the Netherlands including data on patients, tumour, and treatment characteristics, collected from medical records. Linkage of participants

with a negative FIT from the screening database and the cancer registry was established by matching on: initials, birth name, family name, gender, date of birth, postal code, place of birth, and date of death. If an individual with a negative FIT had a CRC registered in the NCR after the date of the FIT analyses in 2014 and before the invitation date of the second screening round, incidence date, and stage (TNM classification) were collected through the registry. To calculate the number of screen-detected CRCs, individuals with a positive FIT in 2014 were similarly linked with the NKR and equivalent data on screen-detected CRCs were collected. All CRCs detected within 6 months after a positive FIT were considered a screen-detected CRC. For staging of CRCs, the 7th edition of the TNM classification was used. Tis (carcinoma in situ) were excluded from the analyses, because these are not invasive cancers. If individuals had more than one CRC diagnosed, for example at two different locations, the CRC with most advanced disease stage was selected for the analyses. The International Classification of Disease for Oncology (ICD-O) was used for coding location and was defined as rectum, rectosigmoid, sigmoid, descending colon, splenic flexure, transverse colon, hepatic flexure, ascending colon, and cecum (C18-C20).[4] Left-sided colorectal cancers included locations from rectosigmoid until descending colon and right-sided colon cancers included locations from splenic flexure to cecum. Appendicael cancers were not considered a CRC in the Dutch CRC screening programme.

Analyses

In all analyses the cut-off of 15 µg Hb/g faeces was referred to as low cut-off, the cut-off of 47 µg Hb/g faeces was referred to as higher cut-off. Proportions of cumulative incidence and sensitivity with 95% confidence intervals (95% CI) were determined by descriptive analyses. The different subgroups (age and sex) were compared using chi-squared test. Because of a substantially different age distribution between the two cut-off groups, we could not use the chi-square test to compare rates by cut-off. Instead, we used multivariable logistic regression analyses to test for statistically significant differences (p <0.05) between the two cut-offs, adjusting for gender and age.

To also facilitate estimates for countries considering different cut-offs than 15 or 47 µg Hb/g faeces, we performed an exploratory analysis to estimate the number of interval CRCs at alternative cut-offs. For every individual we used the absolute concentration of Haemoglobin (Hb) in the sample to determine the numbers of FIT positives and negatives at cut-offs of >0 µg, 10 µg, 20 µg, 40 µg, 60 µg, 80 µg, 100 µg, 120 µg, 140 µg, and 160 µg Hb/g faeces and subsequently we determined how many CRCs would have been missed at those alternative cut-offs. This analysis was based on assumption that all screen-detected CRCs would have become an interval cancer when a cut-off below the measured faecal Hb concentration was applied. Visa versa we assumed that interval CRC after a negative FIT of a certain faecal Hb concentration would have been detected by screening when that concentration was surpassed by the cut-off.

Data analyses were performed using R version 3.5.0.

RESULTS

In the first screening round in 2014, a total of 525,916 individuals had an assessable stool sample, of which 40,942 (7.8%) had a positive FIT and 484,974 (92.2%) had a negative FIT. 127,411 were assessed with the low cut-off; 15,611 (12.3%) had a positive FIT and 111,800 (87.7%) had a negative FIT. 398,505 were assessed with the higher cut-off;25,331 (6.4%) had a positive FIT and 373,174 (93.6%) had a negative FIT. 33,298 (81.3%) of the FIT positives had a colonoscopy follow-up. Among those with a colonoscopy follow-up, 3,210 screen-detected CRCs were diagnosed, 1,102 with a low cut-off and 2,108 with a high cut-off (Table 1). Among those with a negative FIT, 544 interval CRCs were detected: 126 interval CRCs with the low cut-off and 418 interval CRCs with the higher cut-off. Mean age of individuals tested at the low cut-off was 72.0 years and at the higher cut-off was 66.7 years. Median follow-up time between negative FIT and end of interval (invitation subsequent screening round of 24 months for those over 75 years of age) was 730 (IQR 726-730) days. Median follow-up time between negative FIT and date of interval CRC was 469 (IQR 283-618) days. Of all interval CRCs, 188 (34.6%) were detected in the first year after a negative FIT, and 356 (64.4%) were detected in the second year after a negative FIT.

Table 1: Characteristics of the study population by cut-offs (15 and 47 µg Hb/g faeces)

	15 µg	**47 µg**	**Total**
	n (%)	**n (%)**	**n (%)**
Total tested	127,411 (100)	398,505 (100)	525,916 (100)
Gender			
Men	60,936 (47.8)	194,537 (48.8)	255,473 (48.6)
Women	66,475 52.2)	203,968 (51.2)	270,443 (51.4)
Age			
76	54,961 (43.1)	19,256 (4.8)	74,217 (14.1)
75	25,997 (20.4)	52,204 (13.1)	78,201 (14.9)
67	16,103 (12.6)	124,768 (31.1)	140,871 (26.8)
65	28,111 (22.1)	88,340 (22.2)	116,451 (22.1)
63	2,239 (1.8)	86,959 (21.8)	89,198 (17.0)
60	-	26,978 (6.8)	26,978 (5.1)
FIT Negative	111,800 (87.7)	373,174 (93.69)	484,974 (92.2)
FIT Positive*	15,611 (12.3)	25,331 (6.4)	40,942 (7.8)
Screen-detected CRCs	1,102 (0.9)	2,108 (0.5)	3,210 (0.6)
Interval CRCs	126 (0.1)	418 (0.1)	544 (0.1)

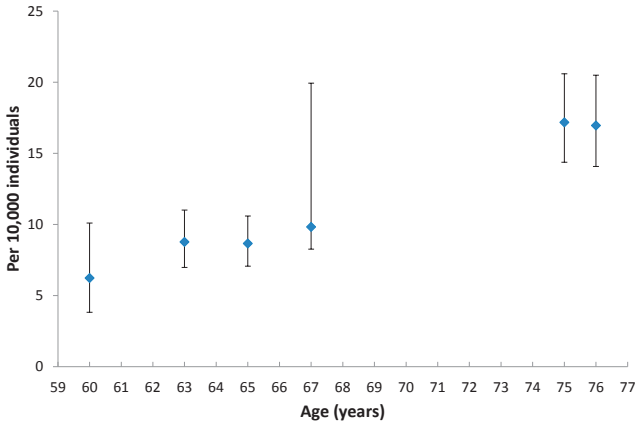
Abbreviations: CRC (colorectal cancer), FIT (faecal immunochemical testing)

*Positive test was defined as a value at or above the cut-off of 15 or 47 µg Hb/g faeces**.**

Cumulative incidence

The cumulative incidence of interval CRC after a negative FIT in the first screening round was 11.2 (95%CI: 10.3-12.2) per 10,000 individuals. The cumulative incidence for men of 12.2 (95%CI: 10.9-13.7) per 10,000 individuals was slightly higher than the cumulative incidence for women of 10.3 (95%CI: 9.2-11.6) per 10,000 individuals, but just not significantly different (p 0.06). Cumulative incidence significantly increased with age (Figure 1; p <0.001). Note, only selected age groups were invited. After adjusting for age differences, the cumulative incidence of interval CRCs was 9.5 per 10,000 individuals at the low cut-off vs 13.8 per 10,000 individuals at the higher cut-off. Multivariable logistic regression analysis showed a significant difference between the two cut-offs, after adjusting for gender and age (p 0.0005).

Figure 1: Cumulative incidence* of interval colorectal cancer with 95% confidence interval after negative FIT



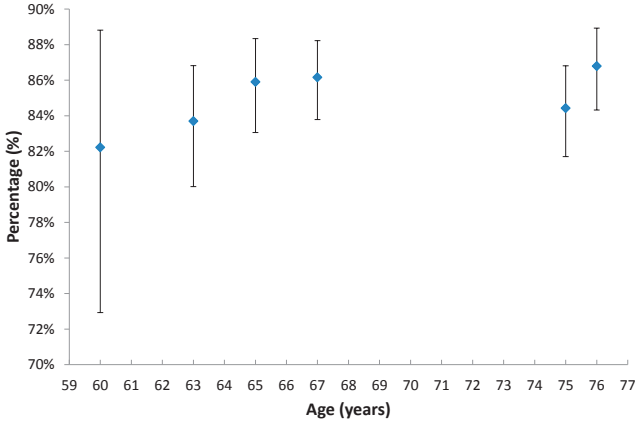
Abbreviations: FIT (faecal immunochemical testing)
*Cumulative incidence is the number of interval CRCs after a negative FIT per 10,000 individuals with a negative FIT.

Sensitivity

Average sensitivity for CRC over both cut-offs in the first screening round was 85.5% (95%CI: 84.3-86.6%). The sensitivity of 87.4% (95%CI: 86.0-88.7%) among men was higher than the sensitivity of 82.6% (95%CI: 80.6-84.5%) among women (p <0.0001). Sensitivity was not significantly different by age (Figure 2; p 0.52). Age-adjusted sensitivity at the low cut-off was 90.5% and 82.9% at the higher cut-off. Multivariable logistic regression analysis showed a significant difference between the two cut-offs, after adjusting for gender and age (p <0.0001).

Exploratory analysis across the full range of relevant cut-offs showed the expected inverse correlation between cut-off and interval CRC rate, with a marked increase in interval CRC rate at high cut-offs (Figure 3). Largest decrease (1.3-0.5%) in positivity rate was observed at low cut-offs (above 0 up to 80 µg Hb/g faeces). Above 80 µg Hb/g faeces approximately

Figure 2: FIT sensitivity with 95% confidence interval for colorectal cancer



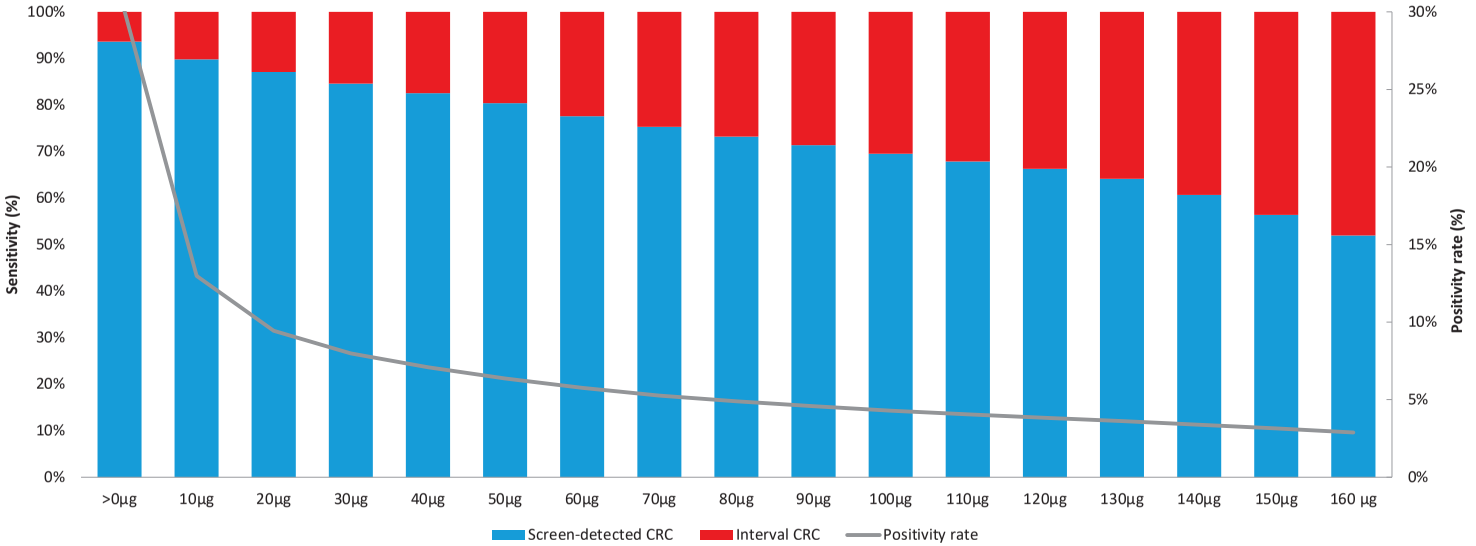
Abbreviations: FIT (faecal immunochemical testing)
*Sensitivity is the number of screen-detected CRCs after a positive FIT divided by the total number of CRCs (screen-detected CRCs and interval CRCs).

0.3% decrease in positivity rate was observed per 10 µg Hb/g faeces increase of FIT cut-off. Contrary, largest decrease in FIT sensitivity for CRC was observed at high cut-offs. FIT sensitivity drops below 70% with cut-offs higher than 90 µg Hb/g faeces, with a sensitivity of only 52.0% at the FIT cut-off 160 µg Hb/g faeces.

Stage distribution and location

A total of 93 (19.8%) stage I interval CRCs were detected, 82 (17.5%) stage II interval, 175 (37.2%) stage III interval CRCs and 120 (25.5%) stage IV interval CRCs. Of 74 (15.7%) interval CRCs stage was unknown. There was no difference between the low cut-off with 75 (63.0%) interval CRCs and the high cut-off with 220 (62.7%) interval CRCs in a late stage (stage III and IV, *p* 0.84) 269 (52.5%) of the interval CRC were located right-sided, 106 (20.5%) left-sided and 141 (27.3%) at the rectum. At the low cut-off a larger proportion of the interval CRCs (119 (57.1%)) was detected right-sided compared to the higher cut-off (397 (50.6%), *p* 0.92)

Figure 3: Positivity rate* and FIT sensitivity† for colorectal cancer at a range of cut-offs



Abbreviations: FIT (faecal immunochemical testing)

*Positivity rate was defined as the number of participants with a test result at or above the cut-off divided by the number of participants with an assessable stool sample.

†Sensitivity is the number of screen-detected CRCs after a positive FIT divided by the total number of CRCs (screen-detected CRCs and interval CRCs).

Table 2: Screen-detected and interval cancers, cumulative incidence and sensitivity using two cut-offs (15 and 47 µg Hb/g faeces)

	Cut-off Hb/g faeces	Negative FITs	Screen-detected CRCs	Interval CRCs	Cumulative incidence*	Sensitivity†
	Total	n	n	n	Per 10,000 individuals (95% CI)	% (95% CI)
All		484,974	3,200	544	11.2 (10.3-12.2)	85.5 (84.3-86.6)
Men		231,138	1,964	282	12.2 (10.9-13.7)	87.4 (86.0-88.7)
Women		253,836	1,246	262	10.3 (9.2-11.6)	82.6 (80.6-84.5)
All	15 µg	111,800	1,102	126	11.3 (9.5-13.4)	89.7 (87.9-91.3)
Age adjusted					*9.5*	*90.5*
Men		52,025	656	73	14.0 (11.1-17.7)	90.0 (87.6-92.0)
Women		59,775	446	53	8.9 (6.8-11.6)	89.4 (86.4-91.8)
All	47 µg	373,174	2,108	418	11.2 (10.2-12.3)	83.5 (82.0-84.9)
Age adjusted					*13.8*	*82.9*
Men		179,113	1,308	209	11.7 (10.2-13.4)	86.2 (84.4-87.9)
Women		194,061	800	209	10.8 (9.4-12.3)	79.3 (76.7-81.7)

Abbreviations: CRC (colorectal cancer), FIT (faecal immunochemical testing)

*Cumulative incidence is the number of interval CRCs after a negative FIT per 10,000 individuals with a negative FIT.

†Sensitivity is the number of screen-detected CRCs after a positive FIT divided by the total number of CRCs (screen-detected CRCs and interval CRCs).

DISCUSSION

In the first screening round of a national FIT-based CRC screening programme, a low incidence of interval CRC in the two years after a negative FIT was observed, irrespective of cut-off. This supports the high FIT sensitivity for CRC. However, the cumulative incidence of interval CRC was higher and sensitivity was lower for individuals tested with the higher cut-off. Older age was associated with a higher interval CRC incidence and FIT sensitivity was lower for women than for men.

We observed a low cumulative incidence of interval CRCs because of a high FIT sensitivity for CRC. Our estimated risk of CRC diagnosis after a negative FIT is approximately 5-fold lower compared to the risk in a similar population before the introduction of CRC screening.[10] The sensitivity in the first screening round for both cut-offs (90.5% and 82.9%) was higher than anticipated (77%), based on the Dutch pilot studies preceding the national programme.[6] There are three potential explanations for this. First, the stability of the buffer of the FIT has been improved. Consequently, higher FIT cut-offs result in similar sensitivity for CRC as lower FIT cut-offs in the past. Second, the median interval between screening rounds was longer in the pilot study (2.4 years) compared to our study (2.0 years).[6] A longer interval could

result in more interval CRCs and therefore may have decreased the sensitivity. However, the third and most important explanation is that we estimated the sensitivity in the first screening round of the national programme, which is a prevalent screening round, while the sensitivity of the pilot study was derived from the total of three screening rounds. In the first screening round relatively more screen-detected CRCs will be detected than in subsequent screening rounds, but the interval CRCs will remain stable, therefore the sensitivity is likely to decrease in subsequent screening rounds. We approximated the FIT sensitivity using screen-detected and interval CRCs, because the real number of CRCs in the population at the moment of screening is unknown. This approximation has three biases. First, sensitivity may be overestimated, because not all missed CRCs will have developed into interval CRCs within two years. This hypothesis is in line with a recent systematic review with all individuals having a colonoscopy follow-up after one-time only FIT, showing a FIT sensitivity for CRC of 71%.[4] Second, some interval CRCs included in the definition of the sensitivity may not have been a missed screen-detected CRC, but still an AA at previous screening. This might lead to an underestimation of the sensitivity. Third, deaths before the end of the interval may have resulted in an overestimation of the sensitivity.

Our estimated FIT sensitivities are at the higher end of those observed in literature.[11-14] However, the Kaiser Permanente group also reported a sensitivity of 85% in the first screening round and then showed a decrease in sensitivity of 6-8% in subsequent screening rounds. Consequently, the sensitivity over four screening rounds was approximately 80%.[11] It is therefore expected that our sensitivity will also decrease in subsequent screening rounds, and will not be that different from the 77% reported in the Dutch pilot studies.[6]

We observed differences between the age-adjusted cumulative incidence and sensitivity between the low and higher FIT cut-off. Despite this difference, the cumulative incidence with the higher cut-off was still low with 13.8 per 10,000 individuals and more than 4 out of 5 CRCs will be detected in the first screening round. Also, with our prior assumption that 12% of the CRC would be missed, we would expect a decrease in sensitivity around 10%. However, the decrease of 7.6% in this study was smaller.[8] The exploratory analysis across the full range of relevant cut-offs showed an increase in interval CRC rate at high FIT cut-offs, which is in line with our main finding. With high FIT cut-offs (>= 160 µg Hb/g faeces) half of the CRCs will probably be missed

The sensitivity for CRC with the higher cut-off was in line with findings of the aforementioned Dutch pilot studies using a cut-off of 10 µg Hb/g faeces and the Kaiser Permanente group using a cut-off of 20 µg Hb/g faeces.[6,11] Again this confirms that the performance of FIT with the old buffer using a low cut-off is comparable to the FIT with the new buffer using a higher cut-off. In a recent systematic review, no difference in sensitivity was observed between different cut-offs, but most included studies used a relatively low cut-off (10-20 µg Hb/g faeces). Nevertheless, the high sensitivity for CRC with a higher cut-off in the current study is promising for many organised programmes using high FIT cut-offs.[15] The results of this study

were based on FOB-Gold screening, but we do expect that they will be generalizable to other FIT brands as a recent study showed comparable performance of FOB-Gold and OC-Sensor.[16] Noteworthy is the difference between the results of the higher cut-offs with FIT in this study compared to sensitivity of guaiac Faecal Occult Blood Testing (gFOBT) of 67.1%.[17]

Our results confirm the higher FIT sensitivity for men than for women.[6,13,14,18] This might raise the question whether different screening strategies for men and women should be applied. However, a decision analysis has shown that risk stratification by gender is currently not effective.[19] We were unable to demonstrate that FIT sensitivity differed by age. This is contrary to other findings suggesting a different sensitivity by age, although the studies presented conflicting results. The increased cumulative incidence for an interval CRC by age can be explained by the higher risk of having a CRC or AA at older age.[13,14,20,21] The stage distribution and location of the interval CRCs were similar for both cut-offs. Interestingly, the stage distribution of interval CRCs is comparable to the stage distribution of clinically detected CRCs, indicating that there probably is no false reassurance after receiving a negative FIT. In contrast, location of the interval CRC is substantially different from that of CRCs detected after symptoms, with many more right-sided interval CRCs, suggesting a lower FIT sensitivity for right-sided CRCs.[22]

The major strength of this study is the opportunity of comparing two FIT cut-offs, applied in the same population within an organised CRC screening programme. We obtained valuable information on the impact of using a higher cut-off. Another strength is the large sample size, using data of a national screening programme. A limitation of the study is that we could not estimate sensitivity for AA. AAs are mostly asymptomatic and therefore not picked up between screenings, and even then not registered at the cancer registry. A recent systematic review showed lower FIT sensitivity for AA than for CRC for one-time testing only.[4,5] However, we expect that missed AA will be detected with repeated FIT in subsequent screening rounds, as AA or an early CRC. Another limitation is that the current conclusions can only be based on the results of selected age groups, due to a phased implementation by birth cohort. Now the full screening programme is implemented, we will assess interval CRCs of all age groups and interval CRCs of subsequent screening rounds.

In conclusion, the incidence of interval CRC after a negative FIT is low. Although FIT sensitivity for declined with a higher cut-off, it remained above 80%.

REFERENCES

1. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;64:1637-49.
2. Lee JK, Liles EG, Bent S, et al. Accuracy of faecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med 2014;160:171.
3. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus faecal immunochemical testing in colorectal-cancer screening. N Engl J Med 2012;366:697-706.
4. Imperiale TF, Gruber RN, Stump TE, et al. Performance Characteristics of Faecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis. Ann Intern Med 2019.
5. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of Benefits, Burden, and Harms of Colorectal Cancer Screening Strategies: Modeling Study for the US Preventive Services Task Force. JAMA 2016;315:2595-609.
6. van der Vlugt M, Grobbee EJ, Bossuyt PMM, et al. Interval Colorectal Cancer Incidence Among Subjects Undergoing Multiple Rounds of Faecal Immunochemical Testing. Gastroenterology 2017;153:439-447 e2.
7. Toes-Zoutendijk E, van Leerdam ME, Dekker E, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Faecal Immunochemical Test Cut-Off Levels. Gastroenterology 2017;152:767-775 e2.
8. Aanpassing uitvoering bevolkingsonderzoek darmkanker in 2014 en 2015. RIVM. 2014. https://www.rivm.nl/sites/default/files/2018-11/Aanpassing%20uitvoering%20bevolkingsonderzoek%20darmkanker%20%20in%202014%20en%202015..pdf. [Cited June 18 2019].
9. Sanduleanu S, le Clercq CM, Dekker E, et al. Definition and taxonomy of interval colorectal cancers: a proposal for standardising nomenclature. Gut 2015;64:1257-67.
10. Selby K, Jensen CD, Lee JK, et al. Influence of Varying Quantitative Faecal Immunochemical Test Positivity Thresholds on Colorectal Cancer Detection: A Community-Based Cohort Study. Ann Intern Med 2018;169:439-447.
11. Jensen CD, Corley DA, Quinn VP, et al. Faecal Immunochemical Test Program Performance Over 4 Rounds of Annual Screening: A Retrospective Cohort Study. Ann Intern Med 2016;164:456-63.
12. Portillo I, Arana-Arri E, Idigoras I, et al. Colorectal and interval cancers of the Colorectal Cancer Screening Program in the Basque Country (Spain). World J Gastroenterol 2017;23:2731-2742.
13. Shin A, Choi KS, Jun JK, et al. Validity of faecal occult blood test in the national cancer screening program, Korea. PLoS One 2013;8:e79292.
14. Wieten E, Schreuders EH, Grobbee EJ, et al. Incidence of faecal occult blood test interval cancers in population-based colorectal cancer screening: a systematic review and meta-analysis. Gut 2018.
15. Digby J, Fraser CG, Carey FA, et al. Can the performance of a quantitative FIT-based colorectal cancer screening programme be enhanced by lowering the threshold and increasing the interval? Gut 2018;67:993-994.
16. de Klerk CM, Wieten E, Lansdorp-Vogelaar I, et al. Performance of two faecal immunochemical tests for the detection of advanced neoplasia at different positivity thresholds: a cross-sectional study of the Dutch national colorectal cancer screening programme. Lancet Gastroenterol Hepatol 2019;4:111-118.

17. 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;58:530-5.
18. Giorgi Rossi P, Carretta E, Mangone L, et al. Incidence of interval cancers in faecal immunochemical test colorectal screening programmes in Italy. J Med Screen 2018;25:32-39.
19. Meulen MPV, Kapidzic A, Leerdam MEV, et al. Do Men and Women Need to Be Screened Differently with Faecal Immunochemical Testing? A Cost-Effectiveness Analysis. Cancer Epidemiol Biomarkers Prev 2017;26:1328-1336.
20. 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;56:1585-9.
21. Wieten E, Schreuders EH, Nieuwenburg SA, et al. Effects of Increasing Screening Age and Faecal Hemoglobin Cutoff Concentrations in a Colorectal Cancer Screening Program. Clin Gastroenterol Hepatol 2016;14:1771-1777.
22. Toes-Zoutendijk E, Kooyker AI, Elferink MA, et al. Stage distribution of screen-detected colorectal cancers in the Netherlands. Gut 2018;67:1745-1746.

6

SOCIOECONOMIC DIFFERENCES IN PARTICIPATION AND DIAGNOSTIC YIELD WITHIN THE DUTCH NATIONAL COLORECTAL CANCER SCREENING PROGRAMME WITH FAECAL IMMUNOCHEMICAL TESTING

Miriam P. van der Meulen, **Esther Toes-Zoutendijk**, Manon C.W. Spaander, Evelien Dekker, Johannes M.G. Bonfrer, Anneke J. van Vuuren, Ernst J. Kuipers, Folkert J. van Kemenade, Marie-Louise F. van Velthuysen, Maarten G.J. Thomeer, Harriët van Veldhuizen, Harry J. de Koning, Iris Lansdorp-Vogelaar, Monique E. van Leerdam, on behalf of the Dutch colorectal cancer screening working group.

Submitted

ABSTRACT

Background

CRC mortality rates are higher for individuals with a lower socioeconomic status (SES). Screening could influence health inequalities. We therefore aimed to investigate SES differences in participation and diagnostic yield of FIT screening.

Methods

All invitees in 2014 and 2015 in the Dutch national CRC screening programme were included in the analyses. We used area SES as a measure for SES and divided invitees into quintiles, with Quintile 1 being the least deprived. Logistic regression analysis was used to compare the participation rate, positivity rate, colonoscopy uptake, positive predictive value (PPV) and detection rate across the SES groups.

Results

Participation to FIT screening was significantly lower for Quintile 5 (67.0%) compared to the other Quintiles (73.0% to 75.1%; adjusted OR quintile 5 versus quintile 1: 0.73, 95%CI: 0.72-0.74), as well as colonoscopy uptake after a positive FIT (adjusted OR 0.73, 95%CI: 0.69-0.77). The detection rate per FIT participant for advanced neoplasia gradually increased from 3.3% in Quintile 1 to 4.0% in Quintile 5 (adjusted OR 1.20, 95CI: 1.16-1.24)). As a result of lower participation, the yield per invitee was similar for Quintile 5 (2.04%) and Quintile 1 (2.00%), both being lower than Quintiles 2 to 4 (2.20%-2.28%).

Conclusions

Screening has the potential to reduce health inequalities in CRC mortality, because of a higher detection in more deprived participants. However, in the Dutch screening programme, this is currently offset by the lower participation in this group.

INTRODUCTION

Colorectal cancer (CRC) is the second most common cause of cancer-related mortality in the Western world.[1] Screening can prevent part of these deaths by early detection and treatment of CRC and its precursor lesions. Therefore, various countries and local initiatives across the world have adopted population-based screening for CRC aiming for equal access to CRC screening for the entire population.[2,3] In Europe, CRC mortality rates are consistently shown to be higher among individuals with a lower socioeconomic status (SES).[4] Since screening can reduce CRC mortality and CRC incidence depending on screening methods and screening uptake, it has the potential to decrease these health inequalities.

However, if the participation to and performance of the screening programme differ across SES groups, screening may fail to reduce or even augment health inequalities. Indeed, several studies demonstrated that lower SES groups had lower participation rates in CRC screening with colonoscopy, guiac faecal occult blood test (gFOBT) and faecal immunochemical test (FIT).[5-11] However, less is known about the participation to subsequent colonoscopy and the performance of a screening programme across SES groups in terms of positivity rate and diagnostic yield. A large study using gFOBT showed that the most deprived individuals had a higher positivity rate and no difference in positive predictive value (PPV).[7] As far as we know, only one small study from the Basque country using FIT showed a similar PPV among SES groups and a higher detection rate in deprived men (but not in women).[12] Because many organised screening programmes across the world have chosen to use FIT, it is important to get more insight into the potential impact of a FIT screening programme on inequalities in health.[3]

Data from the Dutch national CRC screening programme with FIT enabled us to investigate SES differences in participation and diagnostic yield with FIT screening.

METHODS

Dutch CRC screening programme

The Dutch national CRC screening programme using biennial FIT was introduced in 2014 with a gradual roll-out by age within a period of five years. The target population will eventually consist of individuals aged 55 to 75 years. The target population receives a pre-invitation letter by post, followed by an invitation letter by post together with a single FIT sampling device (FOB-Gold, Sentinel, Italy). As a result of the gradual roll-out, in 2014 only individuals aged 63, 65, 67, 75 and 76 years and in 2015 only individuals aged 61, 63, 65, 67, 69 and 75 years were invited. The first half year of 2014, the cut-off level for referral to colonoscopy was 15 µg Hb/g faeces, thereafter, the cut-off level was increased to 47 µg Hb/g faeces because of higher than expected participation rate, positivity rate, and a lower than expected PPV.[13] We

present the data at a cut-off level of 47 µg Hb/g faeces, also for the individuals screened with the lower cut-off level. All data of the screening programme are continuously collected in a national information system of the CRC screening programme (ScreenIT). ScreenIT includes personal details (like gender, date of birth, place of residence, postal code), FIT results, medical details from the pre-colonoscopy intake and colonoscopy results from endoscopy centres and pathology diagnoses from the national pathology registry PALGA. The Dutch screening programme is described in more detail in a previous publication.[13]

Measuring socioeconomic status

Area SES, based on the postal code, was used as a measure for SES. The Dutch postal code consists of four-digits and two letters, of which the four-digit postal code of the invitees' place of residence was used. Scores per four-digit postal code were provided by The Netherlands Institute for Social Research.[14] The provided SES scores per postal code are calculated with a principal components analysis based on income, employment status and educational level.[14] Socioeconomic data of 2014 were used. The scores based on postal codes were divided into quintiles based on the rank of the scores, corrected for the number of individuals (of all ages) living in the postal code areas. The population in the quintiles was calculated with data on the number of inhabitants per age-group in each postal code in 2014.[15] Quintile 1 was the least deprived quintile, with the highest scores (high income, high employment rate, high educational level), while Quintile 5 was the most deprived, with the lowest scores.

Background incidence

Background incidence of CRC across SES groups prior to the introduction of screening was determined as comparator for the yield in FIT participants. All CRC diagnoses from 2008 till 2012 were obtained from the Dutch Cancer Registry (NKR), with the year of diagnosis, the age of the patient at diagnosis and the SES. The SES was determined as described earlier but based on SES scores and population numbers in 2010.

Analysis

National screening programme

Data on the invitees of 2014 and 2015 were collected until 31 March 2016. Outcomes were 1) participation rate of FIT screening, 2) positivity rate of FIT, 3) colonoscopy uptake after a positive FIT, 4) positive predictive value (PPV) for advanced neoplasia (AN, advanced adenomas and CRC combined) and CRC alone, 5) detection rates per participant and 6) yield per invitee of AN and of CRC.

The FIT participation rate was defined as the number of persons returning a stool sample divided by the number of persons invited. Positivity rate was defined as the number of participants with a test result at or above the cut-off level divided by the number of

participants with an assessable stool sample. Cut-off level for a positive test result was 47 ug Hb/g faeces. Positive tests with a result between 15 and 47 ug Hb/g faeces of individuals screened with the lower cut-off level of 15 ug Hb/g faeces were considered as a negative test result and all data collected after the positive test, such as colonoscopy uptake and detected lesions, were not included. The colonoscopy uptake was defined as the number of persons who underwent a colonoscopy divided by the number of persons with a positive FIT. The PPV of AN and CRC was calculated as the number of persons with AN or CRC respectively, divided by the number of persons who underwent a colonoscopy. An advanced adenoma was defined as any adenoma with 1) histology showing ≥25% villous component or 2) high-grade dysplasia or 3) size ≥10 mm. The DR was defined as the number of persons with AN and CRC detected during colonoscopy divided by the number of screened persons with an assessable stool sample (assuming full compliance to colonoscopy). Similarly, the yield per invitee was calculated as the number of persons with AN and CRC detected during colonoscopy divided by the number of invitees. Proportions were determined by descriptive analyses. Logistic regression analysis was performed to estimate odds ratio (OR) of the quintiles on FIT participation rate, positivity rate, colonoscopy uptake, PPV for AN and for CRC and detection rate per invitee for AN and for CRC, adjusted for age and gender. To determine the DR per FIT participant, we performed poststratification (including gender and age) to adjust for the differences in colonoscopy uptake across SES quintiles and assumed full compliance.

Background incidence

Age-standardised incidence rates were calculated by direct standardisation to the European Standard Population (Eurostat 2013).[16] All rates are presented as European age-standardised rates (ESR per 100,000), with 95% confidence intervals (CI). The incidence rate ratio (IRR) was calculated by dividing the ESR of each SES quintile with the corresponding ESR of Quintile 1 (the least deprived quintile), 95% CI were determined.

Sensitivity analysis

In the sensitivity analyses we replicated all analyses with SES divided in deciles instead of quintiles.

The analyses were conducted with R 3.2.3.

RESULTS

Descriptive national screening programme

In 2014 and 2015, 1,882,916 individuals were invited for first round FIT screening, of whom 1,866,060 (99.1%) had an area-based SES score. Quintile 3 contained the largest proportion of invitees (Table 1). Of the invitees with SES score, 49.3% were male, ranging from 48.1% in Quintile 5 to 49.8% in Quintile 2. The invitees of Quintile 5 had a median age of 66.8 years compared with 65.9 years in the total population.

Table 1: Descriptive of the number, age and gender distribution of the invitees in each quintile. Quintile 1 least deprived, Quintile 5 most deprived.

	Overall		Gender male			Age	
Quintile	**n**	**%**	**n**	**%**		**Median**	
1	334,233	17.9	166,013	49.7		65.7	
2	381,344	20.4	189,929	49.8		65.8	
3	403,907	21.6	199,777	49.5		66.0	
4	388,664	20.8	191,341	49.2		66.4	
5	357,912	19.2	172,222	48.1		66.8	
Total	**1,866,060**	**100.0**	**919,282**	**49.3**	***p* <0.001**	**65.9**	***p* <0.001**

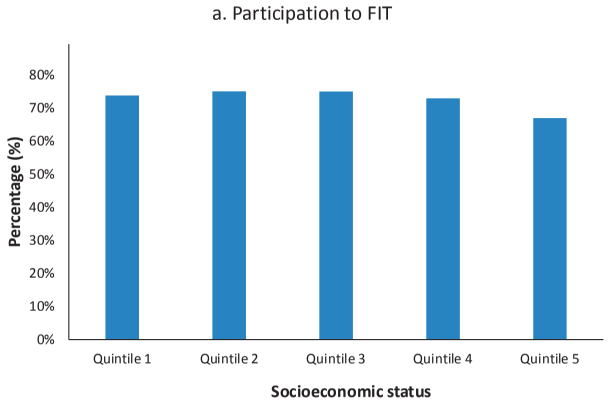
Participation and positivity rate

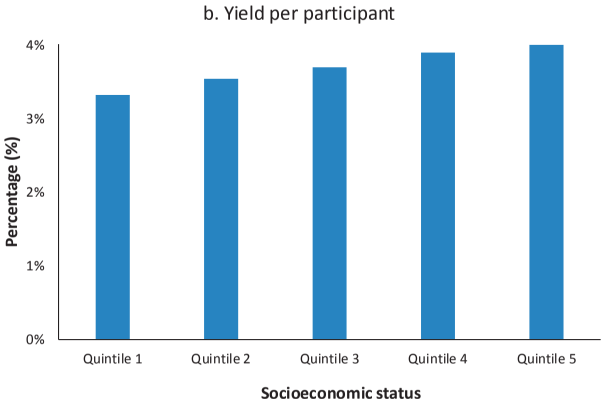
With Quintile 1 as reference, participation to FIT screening was higher in Quintile 2 and 3 (Quintile 1 73.9%, Quintile 2 and 3: 75.1% (Table 2 and Figure 1), but lower in Quintile 4 and Quintile 5, with the lowest participation rate in Quintile 5 (67.0%). Multivariable analysis showed an OR of 0.73 (95% CI: 0.72-0.74) for Quintile 5 compared with Quintile 1. The positivity rate was lowest in Quintile 1 (5.8%) and gradually increased with increasing Quintile. The positivity rate of Quintile 5 (7.2%) had an OR of 1.22 (95%CI: 1.20-1.25) compared to Quintile 1. Colonoscopy uptake after a positive FIT showed a similar pattern as the participation to FIT screening, with the highest uptake in Quintile 2 (82.4%) and significantly lower uptake in Quintile 4 and 5 (80.0% and 75.8% respectively) compared to Quintile 1 (81.3%) (Quintile 5 versus Quintile 1 OR:0.73, 95%CI: 0.69-0.77).

Diagnostic yield

The PPV for AN was highest in Quintile 3 (58.4%) and lowest in Quintile 5 (56.1%). Multivariable analysis showed an OR of 1.06 (95%CI: 1.01-1.12) for Quintile 3 compared with Quintile 1 and an OR of 0.98 (95%CI: 0.93-1.03) for Quintile 5 compared with Quintile 1. The PPV for CRC was also highest in Quintile 3 (9.6%, compared to Quintile 1 adjusted OR: 1.03, 95%CI: 0.95-1.11) and lowest in Quintile 4 (8.5%, compared to Quintile 1 adjusted OR: 0.90, 95%CI: 0.82-0.97; Table 3). The DR for AN in FIT participants was lowest in Quintile 1 (3.33% corrected) and

Figure 1: The participation to FIT (a), detection per participant (b) and yield per invitee (c) of advanced neoplasia for each with the adjusted OR's*.





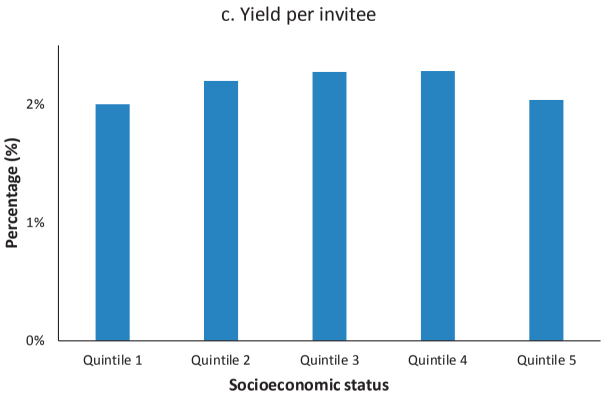


Table 2: The participation to FIT, positivity rate and colonoscopy uptake after a positive FIT in each quintile, with the univariate and multivariable odds ratio (OR) and 95% CI.

Quintile	n	Attendance to FIT	OR*	95% CI	
1	246,858	73.9%	1		*p* <0.001
2	286,527	75.1%	1.07	1.06 - 1.08	
3	303,133	75.1%	1.07	1.06 - 1.08	
4	283,640	73.0%	0.96	0.95 - 0.97	
5	239,945	67.0%	0.73	0.72 - 0.74	
Quintile	**n**	**Positivity rate**	**OR ***	**95% CI**	
1	14,466	5.8%	1		*p* <0.001
2	17,726	6.2%	1.05	1.03 - 1.08	
3	19,235	6.3%	1.08	1.06 - 1.10	
4	19,037	6.7%	1.15	1.12 - 1.17	
5	17,145	7.1%	1.22	1.20 - 1.25	
Quintile	**n**	**Attendance colonoscopy**	**OR***	**95% CI**	
1	11,768	81.3%	1		*p* <0.001
2	14,612	82.4%	1.08	1.02 - 1.14	
3	15,732	81.8%	1.04	0.98 - 1.10	
4	15,234	80.0%	0.93	0.88 - 0.98	
5	12,992	75.8%	0.73	0.69 - 0.77	

* The multivariable OR is corrected for age and gender.

Table 3: The positive predictive value (PPV) of FIT for advanced neoplasia (AN) and colorectal cancer (CRC) in each SES quintile, with the univariate and multivariable odds ratio (OR) and 95% CI.

Quintile	n	PPV AN*	OR**	95% CI	
1	6,689	56.8%	1		*p* <0.001
2	8,388	57.4%	1.02	0.97 - 1.07	
3	9,191	58.4%	1.06	1.01 - 1.12	
4	8,872	58.2%	1.06	1.01 - 1.11	
5	7,295	56.1%	0.98	0.93 - 1.03	
Quintile	**n**	**PPV CRC***	**OR****	**95% CI**	
1	1,103	9.4%	1		*p* <0.01
2	1,376	9.4%	1.00	0.92 - 1.09	
3	1,516	9.6%	1.03	0.95 - 1.11	
4	1,301	8.5%	0.90	0.82 - 0.97	
5	1,165	9.0%	0.94	0.86 - 1.02	

*An advanced adenoma was defined as any adenoma with histology showing ≥25% villous component or high-grade dysplasia or with size ≥10 mm. The PPV was calculated as the number of persons with an advanced adenoma or with a CRC (together called advanced neoplasia (AN) divided by the number of persons who underwent a colonoscopy after a positive FIT.

**The multivariable OR is corrected for age and gender.

gradually increased with higher quintile (Quintile 5: 4.01% corrected; OR: 1.20, 95%CI: 1.16-1.24; Table 4 and Figure 1). The DR for CRC in FIT participants varied between the quintiles and was significantly higher in Quintile 5 with 0.52% (OR: 1.17, 95%CI: 1.08-1.27) compared to Quintile 1. The yield of AN and of CRC in invitees was similar for Quintile 1 and 5, but both Quintiles had significantly lower yield than Quintiles 2 to 4 (Table 4 and Figure 1).

Table 4: The detection rate (DR) per 100 participants uncorrected and corrected for colonoscopy uptake and the yield per 100 invitees of advanced neoplasia (AN) and colorectal cancer (CRC) for each quintile, with the univariate and multivariable odds ratio (OR) and 95% CI.

		PER PARTICIPANT					PER INVITEE			
Quintile	**n**	**DR AN uncorrected***	**DR AN corrected****	**OR*****	**95% CI**		**DR AN**	**OR*****	**95% CI**	
1	6,689	2.71%	3.33%	1		*p*<0.01	2.00%	1		*p* 0.44
2	8,388	2.93%	3.55%	1.07	1.04 - 1.10		2.20%	1.10	1.06 - 1.13	
3	9,191	3.03%	3.70%	1.12	1.09 - 1.15		2.28%	1.13	1.10 - 1.17	
4	8,872	3.13%	3.91%	1.18	1.15 - 1.21		2.28%	1.15	1.11 - 1.19	
5	7,295	3.04%	4.01%	1.21	1.18 - 1.24		2.04%	1.02	0.99 - 1.06	
Quintile	**n**	**DR CRC***	**DR CRC****	**OR*****	**95% CI**		**DR CRC**	**OR*****	**95% CI**	
1	1,103	0.45%	0.55%	1		<0.01	0.33%			*p* 0.07
2	1,376	0.48%	0.58%	1.06	0.98 - 1.15		0.36%	1.09	1.00 - 1.18	
3	1,516	0.50%	0.61%	1.11	1.03 - 1.20		0.38%	1.13	1.04 - 1.22	
4	1,301	0.46%	0.57%	1.04	0.96 - 1.13		0.33%	1.00	0.92 - 1.08	
5	1,165	0.49%	0.64%	1.17	1.08 - 1.27		0.33%	0.97	0.89 - 1.05	

*An advanced adenoma was defined as any adenoma with histology showing ≥25% villous component or high-grade dysplasia or with size ≥10 mm. The detection rate was defined as the number of persons with advanced adenomas or with CRC (together called advanced neoplasia (AN)) detected during colonoscopy divided by the number of screened persons with an assessable stool sample.

**The detection rate was corrected for the differences in colonoscopy uptake compared to Quintile 1.

***The multivariable OR is corrected for age and gender and in the analysis per participant we corrected the DR for non-compliance to colonoscopy using post stratification (assuming full compliance).

Table 5: The number of colorectal cancer cases recorded between 2008 and 2012 and the European age-standardised ratio across the Quintiles of socioeconomic status, and the incidence rate ratio (IRR) of the Quintile compared to the most affluent Quintile (Quintile 1)

Quintile	Incident cases	ESR	95% CI	IRR
1	11,123	456	448 - 465	
2	12,827	467	459 - 475	1.02
3	13,804	466	458 - 474	1.02
4	14,197	471	463 - 478	1.03
5	13,179	462	454 - 470	1.01

Background CRC incidence

In total, 65,130 incident cases of CRC were recorded from 2008 to 2012. The European age-standardised rate was very similar across SES quintiles, varying from 456 per 100,000 in Quintile 1 to 462 per 100,000 in Quintile 5 and was highest in Quintile 4 with 471 per 100,000 (IRR of 1.03; Table 5).

Sensitivity analyses

Using deciles of SES rather than quintiles led to similar patterns in participation, detection and yield, albeit the difference between SES groups was more pronounced. For instance, participation to FIT screening was lowest in Decile 10 with 64.3% compared to 72.6% in Decile 1 (adjusted OR: 0.69, 95%CI: 0.68-0.70). The detection rate per FIT participant for advanced neoplasia gradually increased from 3.2% in Decile 1 to 4.1% in Decile 10 (adjusted OR: 1.28, 95%CI: 1.24-1.33).

DISCUSSION

Our study showed a significantly lower participation to FIT screening and subsequent colonoscopy in case of a positive FIT for individuals in the lowest SES group. The participation was stable for high and moderate SES but decreased for individuals with a low SES. The positivity rate and detection rate of AN gradually and significantly increased with decreasing SES, while the PPV of AN and CRC was quite stable across SES groups.

Even though the participation was lower in Quintile 5, the participation rate of 67.0% in this Quintile was still higher than the desired 65.0% participation rate recommended by the European Union (EU) guidelines for quality assurance.[17] In contrast, the uptake of colonoscopy after a positive FIT was lower than the accepted 85% by the EU guidelines for quality assurance for all quintiles (range 82.4%-75.8%), and was lowest for individuals with a low SES. It is known that the uptake of colonoscopy in case of a positive FIT is higher than registered in the national screening database because some participants opt to have their colonoscopies at centres outside the screening programme. However, we do not expect that individuals with lower SES are more likely to perform the colonoscopy outside the screening programme than those with higher SES and thus do not expect that the observed SES gradient is the result of underreporting.

The SES difference in uptake of colonoscopy can in theory result from a higher prevalence of comorbidity among individuals with lower SES, resulting in exclusion for colonoscopy before or at intake. However, we did not find a difference in ORs for colonoscopy uptake if we corrected for the individuals that were excluded for colonoscopy at intake (data not shown). Another explanation for the association between SES and uptake of colonoscopy is the fact that colonoscopy after a positive FIT is considered standard medical care and is therefore covered by insurance companies. All citizens have an obligatory co-payment for delivered

care during a calendar year ranging between €350 and €850. Therefore, individuals might omit to undergo the procedure or postpone the procedure if this co-payment maximum has not been reached in a given year. This may influence individuals to delay or even forego colonoscopy in order to avoid co-payments, particularly in lower SES.

The positivity rate gradually increased with decreasing SES. Because the PPV of FIT was stable across the SES range, the increase in positivity rate can only be caused by an increase in both true positive (the detection rate) and false positive FIT results. More false positive tests in low SES groups compared to high SES imply that FIT specificity is lower in low SES groups. A possible explanation for the lower specificity could be more comorbidity or anticoagulant use.[18-20]

The increased detection rate in participants with lower SES can either be caused by a higher FIT sensitivity in lower SES for the same reasons as described for specificity or a higher CRC incidence in lower SES. We did not find a difference in CRC incidence by SES quintile for the time period of 2008-2012 (i.e. before the start of the implementation of the national screening program). However, this does not preclude a difference in CRC incidence in those that participate to FIT across SES quintiles. If in lower SES groups individuals with symptoms are more prone to attend screening than individuals without symptoms ("unhealthy screenee bias"), or individuals with an immigrant background are less prone to participate than native Dutch individuals who have a higher CRC incidence, background incidence in the lower SES participants (in contrast with invitees) could be higher than in those with higher SES. Since a previous study observed similar stage distribution of screen-detected CRC across SES quintiles, the first explanation seems unlikely.[21] However, differences in participation between native Dutch and ethnic minorities on the other hand have been previously reported.[22]

Strength of our study is the large sample size and high data completion rate due to the fact that data from different sources were automatically collected in the national screening database ScreenIT, like data on diagnostic yield of the screening programme. Our study also has a limitation; we did not have the personal SES, but based our analysis on the four-digit postal code. These aggregated data on SES may provide an inaccurate representation of the true individual SES. The use of area SES may diffuse results, therefore the observed differences could be more pronounced if linked to personal SES. In theory, there could be a mix of socioeconomic classes in the middle quintiles, but less in quintile 5. In that case the drop in participation might be due to the lack of diffusion in the lowest SES areas.

In other countries with an organised FOBT-based screening programme the smallest socioeconomic difference in participation was 6% (66% for most deprived and 72% for least deprived), while the largest difference was 24% (42% versus 66%).[5] With 67.0% for Quintile 5 versus 75.1% for the middle Quintiles, the difference in participation between SES groups in the Netherlands is at the lower end of this range. The difference between SES groups is also comparable to the differences in the breast cancer screening programme in the Netherlands

(participation rate of 79% for the most deprived up to 87% in the least deprived).[23] The SES differences in yield could also be compared to two other studies. One of those studies used gFOBT instead of FIT and showed a higher positivity rate in higher SES (least deprived), opposite to our findings and a lower PPV for higher SES while we found a stable PPV.[7] A smaller study from the Basque country using FIT was more similar to our results, it showed a similar PPV among SES groups and a higher detection rate in deprived men (but not in women) with an OR of 1.38 (95%CI: 1.23-1.55).[12]

Screening is often argued to increase already existing health inequalities. Based on our data, this is not observed in the Netherlands. Because of the higher yield in lower SES, it even has the potential to decrease health inequalities, however, this is currently offset by the lower participation in lower SES. It is therefore important to know the reasons behind the lower uptake in lower socioeconomic classes. In theory, patient preferences might be different and therefore lead to more individuals not undergoing screening due to a well-informed choice. However, it is more plausible that the lower participation in lower SES is not based on well-informed decision-making, since we previously found that across all quintiles only 12% of non-participants made an informed choice not to participate.[24]

It is difficult to find interventions that decrease the socioeconomic gap in CRC screening. Several interventions have been found to increase overall uptake, such as the involvement of the family doctor. However, most did not reduce the socioeconomic gap or their influence on the socioeconomic gap was not assessed. To date, only two interventions have been demonstrated to reduce the gap, namely targeting specific groups and sending an enhanced reminder letter with a banner that reiterates the screening offer.[25,26] Especially involvement of the family doctor after a positive screening test would be a plausible candidate for decreasing the SES gap in follow-up colonoscopy uptake. However, to recommend this and other specific interventions, further research is needed, also on the underlying reason for non-participation across the socioeconomic groups and to regional and ethnical differences in participation. This research could further clarify how to target groups that are less compliant and/or more at risk for AN and ensure well-informed decision-making.

In conclusion, screening has the potential to reduce existing socioeconomic inequalities in CRC mortality, because of a higher yield in more deprived participants. However, this higher yield is currently offset by the lower participation in this group. Further research is needed into this lower participation to ensure well-informed decision-making.

REFERENCES

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.
2. Gupta S, Sussman DA, Doubeni CA, et al. Challenges and possible solutions to colorectal cancer screening for the underserved. Journal of the National Cancer Institute. 2014;106(4):dju032.
3. Schreuders EH, Ruco A, Rabeneck L et al. Colorectal cancer screening: a global overview of existing programmes. GUT. 2015;64(10):1637-49.
4. Aarts MJ, Lemmens VE, Louwman MW, et al. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. European journal of cancer. 2010;46(15):2681-95.
5. de Klerk CM, Gupta S, Dekker E, et al. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. GUT. 2018;67(4):679-687.
6. Javanparast S, Ward P, Young G, et al. How equitable are colorectal cancer screening programs which include FOBTs? A review of qualitative and quantitative studies. Preventive medicine. 2010;50(4):165-72.
7. Mansouri D, McMillan DC, Grant Y, et al. The impact of age, sex and socioeconomic deprivation on outcomes in a colorectal cancer screening programme. PloS one. 2013;8(6):e66063.
8. Pornet C, Dejardin O, Morlais F, et al. Socioeconomic determinants for compliance to colorectal cancer screening. A multilevel analysis. Journal of epidemiology and community health. 2010;64(4):318-24.
9. Pruitt SL, Shim MJ, Mullen PD, et al, 3rd. Association of area socioeconomic status and breast, cervical, and colorectal cancer screening: a systematic review. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2009;18(10):2579-99.
10. von Wagner C, Good A, Wright D, et al. Inequalities in colorectal cancer screening participation in the first round of the national screening programme in England. British journal of cancer. 2009;101 Suppl 2:S60-3.
11. Wee LE, Koh GC, Chin RT, et al. Socioeconomic factors affecting colorectal, breast and cervical cancer screening in an Asian urban low-income setting at baseline and post-intervention. Preventive medicine. 2012;55(1):61-7.
12. Hurtado JL, Bacigalupe A, Calvo M, et al. Social inequalities in a population based colorectal cancer screening programme in the Basque Country. BMC public health. 2015;15:1021.
13. Toes-Zoutendijk E, van Leerdam ME, Dekker E, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels. Gastroenterology. 2017;152(4):767-75.e2. doi:10.1053/j.gastro.2016.11.022.
14. Statusscores. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores. [Cited November 9 2016].
15. Statistics Netherlands. Statline. http://statline.cbs.nl/StatWeb/default.aspx. Accessed 14 Mar 2012.
16. Eurostat. Revision of the European Standard Population. 2013. http://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f. [Cited October 5 2016].
17. 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 Suppl 3:Se49-64. 18. Brenner H, Tao S, Haug

U. Low-dose aspirin use and performance of immunochemical fecal occult blood tests. Jama. 2010;304(22):2513-20. doi:10.1001/jama.2010.1773.

19. Louwman WJ, Aarts MJ, Houterman S, et al. A 50% higher prevalence of life-shortening chronic conditions among cancer patients with low socioeconomic status. British journal of cancer. 2010;103(11):1742-8.
20. Sasaki K, Kunimoto M, Hirata K. Relationship between the fecal occult blood test and benign anal disorders. Hepato-gastroenterology. 2001;48(38):445-7.
21. Toes-Zoutendijk E, Kooyker AI, Elferink MA, et al. Stage distribution of screen-detected colorectal cancers in the Netherlands. GUT. 2018;67(9):1745-1746.
22. Deutekom M, van Rijn AF, Dekker E, et al. Uptake of faecal occult blood test colorectal cancer screening by different ethnic groups in the Netherlands. European Journal of Public Health. 2009;19(4):400-2.
23. Aarts MJ, Voogd AC, Duijm LE, et al. Socioeconomic inequalities in attending the mass screening for breast cancer in the south of the Netherlands--associations with stage at diagnosis and survival. Breast cancer research and treatment. 2011;128(2):517-25.
24. van Dam L, Korfage IJ, Kuipers EJ, et al. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? European Journal of Cancer. 2013;49(10):2321-30.
25. Decker KM, Demers AA, Nugent Z, et al. Reducing income-related inequities in colorectal cancer screening: lessons learned from a retrospective analysis of organised programme and non-programme screening delivery in Winnipeg, Manitoba. BMJ open. 2016;6(2):e009470.
26. Wardle J, von Wagner C, Kralj-Hans I, et al. Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the English NHS Bowel Cancer Screening Programme (ASCEND): four cluster-randomised controlled trials. Lancet. 2016;387(10020):751-9.

7

QUALITY MONITORING OF A FIT-BASED COLORECTAL CANCER SCREENING PROGRAMME

Esther Toes-Zoutendijk, Johannes M.G. Bonfrer, Christian Ramakers, Marc Thelen, Manon C.W. Spaander, Evelien Dekker, Miriam P. van der Meulen, Maaike Buskermolen, Anneke J. van Vuuren, Ernst J. Kuipers, Folkert J. van Kemenade, Marie-Louise F. van Velthuysen, Maarten G.J. Thomeer, Harriët van Veldhuizen, Marjolein van Ballegooijen, Harry J. de Koning, Monique E. van Leerdam, Iris Lansdorp-Vogelaar.

Clinical Chemistry, 2019

ABSTRACT

Background

Quality assessment is crucial for consistent programme performance of colorectal cancer (CRC) screening programmes using faecal immunochemical test for haemoglobin (FIT). However, literature on the consistency of FIT performance in laboratory medicine was lacking. This study examined the consistency of FIT in testing positive or detecting advanced neoplasia (AN) for different specimen collection devices, lot reagents and laboratories.

Methods

All participants with a FIT sample with a cut-off of 47 µg Hb/g faeces in the Dutch CRC screening programme in 2014 and 2015 were included in the analyses. Multivariable logistic regression analyses were performed to estimate the odds ratios of collection devices, reagents and laboratories, on testing positive or detecting AN and positive predictive value (PPV).

Results

87,519 (6.4%) of the 1,371,169 participants tested positive. Positivity rates and detection rates of AN differed between collection devices and reagents (all p <0.01). In contrast, PPV were not found to vary between collection devices, reagents or laboratories (all p >0.05). Positivity rates showed a small difference for laboratories (p 0.004), but not for detection rates of AN. Size of the population impacted by the deviating positivity rates was small (0.1% of the total tested population).

Conclusions

Variations were observed in positivity and detection rates between collection devices and reagents, but there was no detected variation in PPV. While the overall population-impact of these variations on the screened population is expected to be modest, there is room for improvement.

INTRODUCTION

In recent years many countries have introduced organised screening programmes for colorectal cancer (CRC) using the faecal immunochemical test for haemoglobin (FIT).[1] Such programmes require careful balancing of harms and benefits and adequate quality control.[2] The European Union has therefore developed quality assurance guidelines for CRC screening with the aim to enhance the quality and effectiveness of CRC screening.[3]

In the Netherlands, a national CRC FIT-based screening programme was initiated in 2014. Quality assessment of analytical performance of FIT in the Dutch CRC screening programme consists of three steps: 1) synthetic controls, 2) commutable faeces based controls, and 3) repeated assessment of participant samples with a wide range of FIT results. Recently Fraser et al. (2018) have suggested new analytical performance specifications for FIT, because this quality assessment in the Dutch CRC screening programme has some limitations.[4] Firstly, both internal and external quality assessment are not used for trueness verification purposes as promoted by the 2014 Milan consensus agreement on analytical performance specifications.[5] Secondly, the currently used criteria for accepting reagent and calibrator lots are arbitrary based on expert opinion instead of acceptable impact on predictive value.[6] The expert opinion was dictated as a minimal requirement for the reagent supplier in the tender procedure preceding the implementation of the screening programme and was inherited as an acceptance criterion for laboratory professionals in verifying lot-to-lot variability. However lot-to-lot variability within the acceptance criterion could still result in substantial variation in key screening performance indicators which could not be taken into account when setting the criterion. Thirdly, in its current form the analytical FIT performance specifications within the screening laboratories, which were defined before the start of the Dutch programme in 2014, are not based on their impact on clinical endpoints within the Dutch population-based CRC screening programme.

We therefore examined consistency of FIT over time on clinical endpoints within a national FIT-based CRC screening programme.

METHODS

Screening programme and population

The Dutch national CRC screening programme was initiated in January 2014, with biennial FIT screening for men and women aged 55 to 75 years.[7] The programme has been gradually implemented by birth cohort. Individuals receive an invitation letter and information leaflet with one single FIT specimen collection device (Appendix I). After faecal sampling, FIT samples were sent by postal mail to an assembly point and randomly assigned to one of three central laboratories, where all assessable FITs were analysed. In this study the

laboratories were de-identified by labelling them with a unique letter (LabA, LabB, LabC). If sample return time exceeded 6 days or the FIT specimen collection device was used after the expiration date, individuals with a negative test result were sent a new FIT specimen collection device with the request to resample. All participants with a positive test result were invited for a pre-colonoscopy intake interview and, if no contraindications, referred for colonoscopy. The national screening organisation is responsible for the logistics of the programme and ordering of new FIT specimen collection devices and new reagent lots for the three laboratories.

Faecal immunochemical test

The FOB-Gold (Sentinel Diagnostics SpA, Milan, Italy), an automated quantitative FIT, was used. In this reagent, polyclonal anti-human haemoglobin (Hb) antibodies coated on polystyrene beads, bind human haemoglobin thereby triggering an agglutination reaction that can be quantified by turbidity. Samples were pre-analytically processed using an Impeco track. Turbidity analysis was performed by a track connected JCA-BM6010 Biomajesty clinical chemistry analyser (Jeol). The FIT specimen collection devices contain a buffer solution (1.7 mL) and a green screw-cap with faecal collector. All specimen collection devices were labelled with a unique bar code. Different batches with FIT collection devices, all FOB-Gold (Sentinel Diagnostics SpA) were labelled with a three letter code; these are the first three letters of the unique bar code on the FIT specimen collection device (AAA, AAC, AAD, AAE). During the pilot phase at the end of 2013, preceding the start of the national programme, specimen collection device AAB was used. This data was not included in the study. The FIT manufacturer claims Hb stability in the FIT specimen collection device for 14 days at 2-8 °C or 7 days at 15-30 °C. A sample taken from the FIT specimen collection device is mixed with latex reagent. An immunological latex reagent is used for the determination of haemoglobin, added and mixed with the buffer solution in the FIT specimen collection device. Different reagent lots were also labelled with a unique lot number (1-6). The cut-off for a positive test result was initially set at 15 µg Hb/g faeces. As the programme performed differently than the predefined programme indicators, the cut-off was increased to 47 µg Hb/g faeces in July 2014.[7]

Quality control in the Dutch programme

Pre-analytical aspects and analytical performance of the FIT analysis are described in the FITTER checklist (Appendix II). Daily controls in the participating laboratories, supervised by the Dutch Foundation for Quality Assessment in Medical Laboratories (SKML) consist of three groups of controls: 1) Synthetic controls provided by the IVD manufacturer, 2) Commutable faeces based controls with addition of known amounts of human haemoglobin provided by EQAS organiser SKML, 3) The percentage of patients above and below certain predefined values in the participating laboratories. Laboratories determine analytical performance by

judging their own control performance on a daily basis and by judging the comparison of the three laboratories on weekly basis supported by SKML. SKML also provides commutable faeces based external quality assessment with known amounts of haemoglobin that are blinded to the laboratories. These samples are provided in 6 rounds of 15 samples each with values across the measurement range. Laboratories measure external quality assessment on a weekly basis and receive reports every two months. Prior to reagent acceptance for measurement by the participants, all reagent and calibrator lot combinations are compared to the previous reagent and calibrator of the same product using internal control of Sentinel, that of SKML, and 40 participants' samples. Acceptance criteria state that all results must be within ±7.5% of the overall-all-lot mean. This approach aims to prevent a worst-case difference between lot differences of larger than 15%. During the pilot phase at the end of 2013, preceding the start of the national programme, only one reagent lot was not accepted because it did not meet these criteria. The results in this study were only measured with accepted reagent lots.

Data collection

This study included all participants invited in 2014 and 2015. For the purpose of these analyses, only data of participants with an assessable faecal sample were analysed. Data were obtained from the national CRC screening database (ScreenIT). Individuals who objected to exchange of data were not included in the analyses. ScreenIT includes date of invitation, date of faecal sampling, date of analysis, concentration µg Hb/g faeces and colonoscopy and pathology reports. Data were collected until April 24, 2017. Data on reagent and time to expiry date of collection devices were collected through the regional screening organisations. Data on the ambient temperature were collected from the Royal Netherlands Meteorological Institute (KNMI).

Measures and definitions

The outcomes of interest were FIT positivity rate, detection rate of advanced neoplasia (AN), and PPV. ANs are considered relevant findings within the Dutch CRC screening programme and consist of CRCs and advanced adenomas (AAs). An AA is defined as any adenoma with histology showing >=25% villous component or adenoma with high-grade dysplasia or with size >=10 mm. FIT positivity rate was calculated as the number of individuals with a test result at or above the cut-off divided by number of participants with an assessable FIT. For the purpose of these analyses, all FITs were only considered positive at a cut-off of 47 µg Hb/g faeces, regardless of the 15 or 47 µg Hb/g cut-off applied within the national screening programme. The same applies to colonoscopy results; these outcomes were also only included in the analyses if the FIT result was above a cut-off of 47 µg Hb/g faeces. Detection rate was considered as the number of individuals with AN detected during colonoscopy

divided by the number of screened individuals. PPV was defined as the number of individuals with AN among individuals with a positive FIT who underwent a diagnostic colonoscopy.

Outcome variables were FIT specimen collection devices, reagent lots, laboratories, sex, age, ambient temperature, sample return time and time to expiry date of the test. Age was determined at the self-sampling date. Ambient temperature was assessed at the date of analysis of the faeces sample minus one day, based on the assumption that the faecal sample was outdoors during the transportation phase. As a reference place for the average ambient temperature we used a geographically central location in the Netherlands (De Bilt). Sample return time was defined as the interval in weekdays between self-sampling date and analysis date. Negative values of sample return time were coded as missing, as these were data entry errors. As only data of samples with a sample return time ≤6 days with a positive test result were analysed and not those with a negative test result, for the purpose of this analysis individuals with values >6 days (positive and negative) were removed. The time to expiry date of the test was the interval between the date of analysis and the expiry date indicated on the FIT specimen collection device. Individuals with negative values were coded as missing and samples exceeding the expiry date were removed from the analysis.

Statistical analysis

Descriptive analysis were conducted to determine means and proportions of baseline characteristics. The Pearson chi-square test was used for the comparison of dichotomous or categorical variables and the *t*-test was used for the comparison of continuous variables. Multivariable logistic regression analyses were performed to estimate the odds ratios (OR) for FIT specimen collection device, reagent lots and laboratories, for a positive test result, detection of AN and PPV. Outcomes were adjusted for sex, age, ambient temperature, sample return time, and time to expiry date of test. Continuous variables were modelled with splines, using 3 knots, except for sample return time. This variable had one spike (majority having same value of 1 day), which was therefore added as a categorical variable. Overall significance was tested with ANOVA.

Two uncertainty analyses were performed. The first analysis assessed whether the results of the multivariable logistic regression analyses changed when adding a combined variable of FIT specimen collection device and reagent lots in the multivariable logistic regression model. The second analysis was carried out to assess whether results changed when the multivariable logistic regression analyses were stratified by age. Because the programme was gradually implemented by year of birth, age was highly correlated with collection devices and reagent. The age groups 63 and 67 years were chosen for evaluation of the age effect as these age groups comprised the largest number of individuals in each subgroup of categorical variables. All statistical tests were two-tailed and P <0.05 was considered statistically significant. Statistical analyses were performed with the R statistical package version 3.2.3.

RESULTS

Study population characteristics

A total of 1,372,020 first-round participants had an assessable FIT. In 851 (0.1%) individuals the sample return time exceeded 6 days. They were excluded from the analysis, leaving 1,371,169 participants for analyses. Of these 663,884 (48.4%) were men. The majority of participants, 1,313,052 (96.3%) returned their faecal sample within two days after sampling. Of all participants with an assessable FIT, 87,519 (6.4%) tested positive, 71,931 (82.2%) of these individuals underwent colonoscopy. Results of colonoscopy and/or pathology were available for 71,753 (99.8%) individuals. Of those, 6,636 (9.2%) were diagnosed with CRC and 34,803 (48.5%) with AA, with a PPV for AN of 57.8%. All baseline characteristics of participants differed between FIT positives and FIT negatives, except time to expiry of the test (Table 1).

Multivariable logistic analyses

Multivariable logistic regression analysis showed that participants tested with FIT specimen collection device AAE are less likely to test positive than collection device AAA (OR:0.82, 95%CI: 0.73-0.92; Table 2). There were also differences in the probability of testing positive between individuals analysed with different reagent lots (OR ranging from 1.11 to 1.31, p <0.001; Table 2). Individuals that had their FIT analysed in laboratory B and laboratory C tested positive more often (p 0.004; Table 2), however the effect size was very small (LabB OR:1.03, 95%CI: 1.01-1.05; LabC OR:1.02, 95%CI: 1.01-1.04).

The detection rate of AN showed a similar pattern as the positivity rate with respect to FIT specimen collection device and reagents lots (device p <0.001; reagent lots p 0.004; Table 3). Detection rate of AN was especially lower for participants tested with device AAE (OR:0.77, 95%CI: 0.65-0.91) and higher for participants tested with reagent lot 6 (OR:1.23, 95%CI: 1.07-1.41). No difference was observed between the three laboratories in having a diagnosis with AN (p 0.37).

PPV for AN did not significantly differ for any of the variables of interest (device p 0.70, reagent lot p 0.96, laboratory p 0.23; Table 4).

Uncertainty Analyses

When adding a combined variable of batch FIT specimen collection devices and reagent lot into the multivariable logistic regression analysis for testing positive, an association between this combined variable and testing positive remained. Remarkable was that a batch in combination with the latest added reagent lot (highest number) resulted more often in a positive FIT test result: AAA3, AAC5 and AAD6 (Appendix III). Largest deviating positivity rates were observed in the combination AAA3 (OR:1.31, 95%CI: 1.02-1.69) and AAC5 (OR:1.51, 95%CI: 1.05-2.18); however the population tested with these two combinations was very

Table 1: Baseline characteristics of participants with an assessable faeces sample

	FIT negatives	FIT positives*	Total	*p* value
Total (n, %)	1.283.650 (93.6)	87.519 (6.4)	1.371.169 (100)	<0.001
Sex (n, %)				
Men	611,125 (92.1)	52,759 (7.9)	663,884 (100)	<0.001
Women	672,525 (95.1)	34,760 (4.9)	707,285 (100)	
Age (mean, sd)				
Year	66.9 (4.3)	67.7 (4.6)	67.0 (4.4)	<0.001
Device (n, %)				
AAA	304,754 (93.3)	22,018 (6.7)	326,772 (100)	<0.001
AAC	308,413 (93.8)	20,292 (6.2)	328,705 (100)	
AAD	575,421 (93.7)	38,811 (6.3)	614,232 (100)	
AAE	95,062 (93.7)	6,398 (6.3)	101,460 (100)	
Reagent lot (n, %)				
1	249,787 (93.3)	17,937 (6.7)	267,724 (100)	<0.001
2	232,445 (93.7)	15,557 (6.3)	248,002 (100)	
3	235,090 (93.7)	15,770 (6.3)	250,860 (100)	
4	233,284 (93.9)	15,224 (6.1)	248,508 (100)	
5	158,091 (93.8)	10,523 (6.2)	168,614 (100)	
6	174,953 (93.3)	12,508 (6.7)	187,461 (100)	
Laboratory (n, %)				
LabA	414,232 (93.5)	28,651 (6.5)	442,883 (100)	<0.001
LabB	423,476 (93.5)	29,300 (6.5)	452,776 (100)	
LabC	445,942 (93.8)	29,568 (6.2)	475,510 (100)	
Ambient temperature (mean, sd)				
degrees Celsius	10.8 (5.3)	10.7 (5.3)	10.8 (5.3)	0.02
Sample return time (n, %)				
1 day	1,042,322 (94.1)	65,167 (5.9)	1,107,489 (100)	<0.001
2 days	192,919 (93.8)	12,644 (6.2)	205,563 (100)	
3 days	30,823 (93.7)	2,077 (6.3)	32,900 (100)	
4 days	10,202 (93.6)	700 (6.4)	10,902 (100)	
5 days	3,852 (93.7)	258 (6.3)	4,110 (100)	
6 days	2,776 (93.8)	182 (6.2)	2,958 (100)	
Time to expiry of device (mean, sd)				
Days	293.4 (83.0)	293.3 (84.1)	293.4 (83.0)	0.69

Abbreviations: FIT (faecal immunochemical test for haemoglobin).
*FITs were considered positive at a cut-off of 47 µg Hb/g faeces.

small (0.1%). Additionally, the combination of AAD6 and AAA1 resulted in deviating positivity rates, however the effect size was smaller (OR:1.14, 95%CI: 1.08-1.20 and OR:0.92, 95%CI: 0.88-0.95, respectively). Although these combinations showed smaller ORs, they affected a larger group of individuals (24.8%). The remaining combinations were not significantly different.

Stratifying the multivariable models for two age groups for testing positive resulted in similar effects sizes as the full model, although some differences were no longer statistically significant because of the longer sample size.

Table 2: FIT* test results by specimen collection device, reagent, laboratory and multivariable logistic analysis**

	Positivity rate (95% CI)	OR (95% CI)	*p* value
Device			
AAA	6.7 (6.7-6.8)	REF	<0.001
AAC	6.2 (6.1-6.3)	0.98 (0.94-1.03)	
AAD	6.3 (6.3-6.4)	0.94 (0.87-1.02)	
AAE	6.3 (6.2-6.5)	0.82 (0.73-0.92)	
Reagent lot			
1	6.7 (6.6-6.8)	REF	
2	6.3 (6.2-6.4)	1.11 (1.06-1.16)	<0.001
3	6.3 (6.2-6.4)	1.14 (1.08-1.21)	
4	6.1 (6.0-6.2)	1.14 (1.06-1.22)	
5	6.2 (6.1-6.4)	1.13 (1.04-1.24)	
6	6.7 (6.6-6.8)	1.31 (1.19-1.44)	
Laboratory			
LabA	6.5 (6.4-6.5)	REF	0.004
LabB	6.5 (6.4-6.5)	1.03 (1.01-1.05)	
LabC	6.2 (6.2-6.3)	1.02 (1.01-1.04)	

Abbreviations: FIT (faecal immunochemical test for haemoglobin), OR (Odds ratio), CI (confidence interval).
*FITs were considered positive at a cut-off of 47 µg Hb/g faeces.
** Multivariable OR were corrected for sex, age, ambient temperature, sample return time and time to expiry of collection device

Table 3: Detection rates of advanced neoplasia* by specimen collection device, reagent, laboratory and multivariable logistic analysis **

	Number of participants	Number of individuals with advanced neoplasia (detection rate (95% CI))	OR (95% CI)	*p* value
Device				
AAA	326,772	10,624 (3.3 (3.2-3.3))	REF	<0.001
AAC	328,705	9,529 (2.9 (2.8-3.0))	0.96 (0.90-1.02)	
AAD	614,232	18,500 (3.0 (3.0-3.1))	0.92 (0.82-1.02)	
AAE	101,460	2,786 (2.7 (2.6-2.8))	0.77 (0.65-0.91)	
Reagent lot				
1	267,724	8,644 (3.2 (3.2-3.3))	REF	0.004
2	248,002	7,392 (3.0 (2.9-3.0))	1.07 (1.01-1.14)	
3	250,860	7,403 (3.0 (2.9-3.0))	1.10 (1.01-1.19)	
4	248,508	7,310 (2.9 (2.9-3.0))	1.10 (1.00-1.22)	
5	168,614	5,074 (3.0 (2.9-3.1))	1.11 (0.98-1.25)	
6	187,461	5,616 (3.0 (2.9-3.1))	1.23 (1.07-1.41)	
Laboratory				
LabA	442,883	13,530 (3.1 (3.0-3.1))	REF	0.37
LabB	452,776	13,801 (3.0 (3.0-3.1))	0.99 (0.97-1.02)	
LabC	475,510	14,108 (3.0 (2.9-3.0))	0.98 (0.96-1.01)	

Abbreviations: OR (Odds ratio), CI (confidence interval).

*Advanced neoplasia was defined as CRCs and advanced adenomas (AA). AA is defined as any adenoma with histology showing ≥25% villous component or high-grade dysplasia or adenoma with size ≥ 10 mm.

** Multivariable OR were corrected for sex, age, ambient temperature, sample return time and time to expiry of collection device

DISCUSSION

In a well-organised FIT-based screening programme with strong focus on quality assurance, FIT positivity rates varied by FIT specimen collection devices, reagent lot, and laboratories as well as detection of AN for FIT specimen collection devices and reagent lot. These effects remained after multivariable correction for sex, age, ambient temperature, sample return time and time to expiry of collection device. The PPV for AN were not found to differ between FIT specimen collection devices, reagent lots and laboratories. The small difference between the three laboratories responsible for the analyses in the national Dutch CRC screening programme is considered clinically irrelevant.

The observed differences in this study were surprising and unexpected, as we currently have a thorough quality assessment programme in the Netherlands. With every reagent lot change quality assessments are in place and quality assessments are regularly carried out in the laboratories. On the other hand, the results are not that surprising, considering the

many uncontrolled factors that can influence the quality of FIT specimen collection devices and reagent: variations in composition of the tube material, buffer, brush, stick, antibodies and so on. Remarkable was that the largest observed differences were predominantly observed in FIT specimen collection devices that were analysed with a newer reagent lot, as shown in the uncertainty analysis. These differences were seen in individuals that were sent a FIT during a certain reagent, but who waited a considerable time before returning their FIT during which a new reagent lot was introduced. This difference may be the result of selection bias, if individuals that wait longer to return their FIT are a selected group of individuals with more CRC or AA but also more comorbidities. An alternative explanation might be that the manufacturer calibrates a new reagent lot on the buffer of new specimen collection devices, and not on the old devices. Fortunately, the clinical impact of the difference in

Table 4: PPV for advanced neoplasia* at colonoscopy (PPV) by specimen collection device, reagent, laboratory and multivariable logistic analysis **

	Number of positive FITs	Number of individuals with colonoscopy (participation rate (95% CI))	Number of individuals with advanced neoplasia (PPV (95% CI))***	OR (95% CI)	*p* value
Device					
AAA	22,018	17,959 (81.6 (81.0-82.1)	10,624 (59.4 (58.6-60.1))	REF	0.70
AAC	20,292	16,658 (82.1 (81.6-82.6))	9,529 (57.3 (56.5-58.0))	0.99 (0.89-1.09)	
AAD	38,811	32,132 (82.8 (82.4-83.2))	18,500 (57.7 (57.2-58.3))	0.97 (0.82-1.14)	
AAE	6,398	5,182 (81.0 (80.0-81.9))	2,786 (53.9 (52.5-55.2))	0.90 (0.69-1.16)	
Reagent lot					
1	17,937	14,648 (81.7 (81.1-82.2))	8,644 (59.2 (58.4-60.0))	REF	0.96
2	15,557	12,754 (82.0 (81.4-82.6))	7,392 (58.0 (57.2-58.9))	0.97 (0.89-1.06)	
3	15,770	12,997 (82.4 (81.8-83.0))	7,403 (57.1 (56.2-57.9))	0.96 (0.85-1.09)	
4	15,224	12,702 (83.7 (83.1-84.2))	7,310 (57.5 (56.7-58.4))	0.95 (0.81-1.11)	
5	10,523	8,688 (82.7 (82.0-83.4))	5,074 (58.4 (57.4-59.4))	0.97 (0.80-1.17)	
6	12,508	10,065 (80.7 (80.0-81.4))	5,616 (55.8 (54.8-56.8))	0.95 (0.77-1.17)	
Laboratory					
LabA	28,651	23,607 (82.4 (81.9-82.8))	13,530 (57.5 (56.8-58.1))	REF	0.23
LabB	29,300	24,028 (82.0 (81.6-82.4))	13,801 (57.6 (56.9-58.2))	1.00 (0.96-1.04)	
LabC	29,568	24,296 (82.2 (81.7-82.6))	14,108 (58.2 (57.6-58.8))	1.03 (0.99-1.07)	

Abbreviations: OR (Odds ratio), CI (confidence interval), PPV (Positive Predictive Value), FIT (faecal immunochemical test for haemoglobin).

*Advanced neoplasia was defined as CRCs and advanced adenomas (AA). AA is defined as any adenoma with histology showing ≥25% villous component or high-grade dysplasia or adenoma with size ≥ 10 mm.

** Multivariable ORs were corrected for sex, age, ambient temperature, sample return time and time to expiry of collection device**.**

***In total, 71,931 (82.2%) individuals underwent colonoscopy. Results of colonoscopy and/or pathology were missing for 178 (0.2%) individuals. Those individuals were not included in the calculation of the PPV.

positivity rate on population level is very small, as only 0.1% of the population tested had these significantly higher odds ratios of 1.3 or higher.

To our knowledge, no previous studies have examined the performance of the FIT test, using different FIT specimen collection devices, reagent lots and laboratories. Our findings for the confounding factors were consistent with literature (data not shown).[8,9] Therefore we expect that the findings of our study are not unique for FOB-Gold or the Netherlands, but are generally applicable for FIT-based screening programmes throughout the world. Other studies have shown instability of the positivity rate of FIT-based screening programmes, which even resulted in temporary suspension of some programmes.[10,11] The variation reported here is not of the same magnitude and requires other measures such as the development of a regular monitoring system. This emphasises an important caveat when publishing data on FIT screening, as continuous improvements to FIT systems by manufacturers affect the relevance of the results over time. Although results cannot directly be used by other programmes, these results highlight the importance of careful and continuous monitoring of the FIT results.

A key strength of the present study is that it is the first to report on consistency in performance of FIT at population level. Another important strength is the large population-based design, which made it possible to examine the performance of FIT and analytical performance on a large scale. There are also some noteworthy limitations. First, there might be a correlation between specimen collection devices and reagent, as both were gradually introduced consecutively. However, we tested this correlation with the sensitivity analysis by looking at a combined variable, and obtained similar results to the main analysis. A correlation is also possible between specimen collection device, reagent and season, but we tried to diminish the impact by correcting for ambient temperature and using data of two screening years. Another limitation is the use of positivity rate as performance indicator instead of distribution of faecal haemoglobin concentration.[12] However, in this study we focused on the real programme performance, which was reflected by the positivity rate. The largest shortcoming is intrinsic to the concept of screening; we could not assess the negative predictive value, as individuals with negative FIT results were not assessed by endoscopy.

Although we hypothesised the cause of the observed difference to be related to a specific group of individuals or combination of reagent and buffer, we also found odds ratios significantly different from 1 for a different and larger group of individuals: 15% higher odds for the combination of device AAD and reagent lot 6 and 8% lower odds for the first combination used, device AAA with reagent lot 1. These effect sizes are much smaller and therefore the clinical impact will be less important. They do, however, affect a considerably larger number of individuals (almost 25% of the tested population). These results therefore emphasise the importance of the need for standardisation of quality control of FIT performance. Currently no standard exists for acceptable range of variation on the outcomes of a FIT-based CRC screening programme. Although SKML EQA samples have

target values assigned by weighed in addition of human haemoglobin and therefore could be used for trueness verification, this is not the current practice due to debate on the value assignment as a lack of international standardisation of the method for value assignment in FIT testing and agreement of performance specifications. Although manufacturers assess the quality of the FIT, they use small sample sizes and data is not publicly available. Again, this highlights the necessity of standardising quality assessment. To define this standard, the International Federation of Clinical Chemistry and Laboratory Medicine has set up a Faecal Immunochemical Testing working group (WG-FIT).[13] Additionally, more information is needed on the long-term impact of this observed variation of FIT. Therefore we are currently performing a decision analysis to define a range in which variation is acceptable without affecting long-term effect of screening (by means of mortality reduction and life years gained). Furthermore, evaluation of the effect of analytical performance on clinical outcomes is planned following the approach of the European Federation of Clinical Chemistry and Laboratory Medicine.[13] In addition to standardising quality assessment, a three-step plan must be set up: 1) controls in the laboratory itself, 2) regular monitoring of FIT performance, and 3) evaluation of FIT performance at population level. All steps are currently followed in the Netherlands, but the lack of international standardisation of methods, value assignment and acceptance criteria hamper correlating analytical performance and screening efficiency. Another possible approach to reduce the observed differences could be synchronisation of specimen collection device and reagent lot changeover, however this is logistically difficult to implement in the national screening programme.

Conclusions

Test positivity of FIT and detection rates of AN differed between FIT specimen collection devices and reagent lots, however no variation in PPV was observed. Clinically the programme is performing well, but there is room for improvement. Acceptable ranges of variation are lacking. These results can be used as input for the international initiative for standardising FIT quality assessment and for improving a regular monitoring system to reduce the impact of test variation on detection of AN.

REFERENCES

1. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmemes. Gut 2015;64:1637-49.
2. 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;8:317-9.
3. Halloran SP, Launoy G, Zappa M, International Agency for Research on Cancer. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy 2012;44 Suppl 3:SE65-87.
4. Fraser CG, Benton SC. Detection capability of quantitative faecal immunochemical tests for haemoglobin (FIT) and reporting of low faecal haemoglobin concentrations. Clin Chem Lab Med 2019;57(5):611-616.
5. Kallner A, McQueen M, Heuck C. The Stockholm Consensus Conference on quality specifications in laboratory medicine, 25-26 April 1999. Scand J Clin Lab Invest 1999;59:475-6.
6. McDonald P, Anderson CM, Fraser CG. Acceptance quality checks for qualitative faecal immunochemical tests ensure screening programme consistency. Int J Cancer 2011;128:247-8.
7. Toes-Zoutendijk E, van Leerdam ME, Dekker E, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Programme and Optimization by Altering Faecal Immunochemical Test Cut-Off Levels. Gastroenterology 2017;152:767-75 e2.
8. Dancourt V, Hamza S, Manfredi S, et al. Influence of sample return time and ambient temperature on the performance of an immunochemical faecal occult blood test with a new buffer for colorectal cancer screening. Eur J Cancer Prev 2016;25:109-14.
9. Grazzini G, Ventura L, Rubeca T, et al. Impact of a new sampling buffer on faecal haemoglobin stability in a colorectal cancer screening programmeme by the faecal immunochemical test. Eur J Cancer Prev 2016; 26:285-291.
10. Australian Institute of Health and Welfare Canberra. National Bowel Cancer Screening program. https://www.aihw.gov.au/getmedia/832d1b18-7fbf-4594-8673-12251fcc1930/13050.pdf.aspx?inline=true. [Cited February 6 2018].
11. Provincial Health Services Authority. What is the FIT? http://www.bccancer.bc.ca/screening/colon/get-screened/what-is-the-fit. [Cited February 6 2018].
12. Fraser CG. Assessment of faecal haemoglobin concentration distributions is vital for faecal immunochemical test (FIT)-based colorectal cancer screening programmes. J Med Screen 2016;23:52-3.
13. Sandberg S, Fraser C, Horvath A, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2015; 53:833–5.

APPENDIX I: INFORMATION LEAFLET WITH INSTRUCTIONS FOR USE OF THE FAECAL IMMUNOCHEMICAL TEST KIT

bevolkingsonderzoek

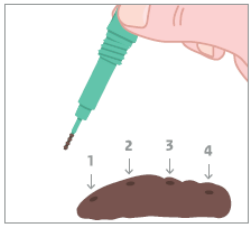
Instructions for use

These instructions include a step-by-step description of how to perform the self-sampling test properly.

Collect a tiny amount of stool material, using the ribbed stool sampling stick. If too much stool material is taken, the results of the test will be unusable.

There are a few things you should know in advance

- You can only use the sampling tube once.
- Place the sampling tube and the reply form in the return envelope.
- The sampling tube is intended only for the individual who received the invitation for this test.
- Before use, the sampling tube should not be stored at temperatures below 2°C or above 30°C. You can store the tube in the refrigerator.
- It is important to limit the time the tube is in transit. You should send the envelope on a day that the mail box in your area is emptied. You should not post the sample on the weekend, or right before a public holiday.



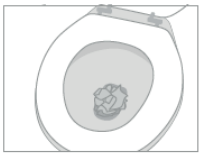
Bowel cancer screening programme | 1

Important

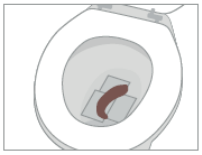
- Only a tiny amount of stool material is needed to perform the test properly.
- Your stool should not come into contact with urine or water.
- You must remove only the green cap.

How to carry out the test?

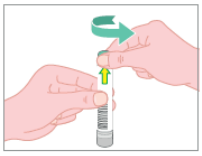
1 The test is best done on a Monday, Tuesdays, Wednesday or Thursday.
Important: your stool should *not* come into contact with urine or with the water in the toilet. You should, therefore, line the lavatory bowl with several sheets of toilet paper.



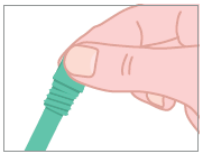
2 You can also use a wad of toilet paper. For more tips about how to collect a stool sample, go to www.bevolkingsonderzoekdarmkanker.nl



3 Go to the toilet and make sure that your stool falls onto the paper.

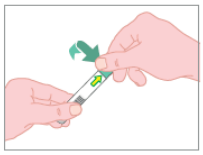


4 Twist the **green cap** to open the tube. Attached to the cap is a stool sampling stick with a ribbed section at its tip.

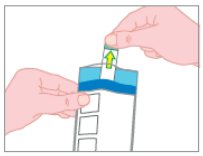


5 Dip the ribbed section of the stool sampling stick into four different parts of your stool. No more than a **tiny amount of stool** should stick to the ribbed section. Too much stool material makes the tube unusable.

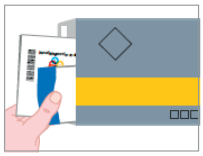
6 Even if only a very little stool material sticks to the stool sampling stick, this is sufficient to carry out the test effectively.
Please note: if too much stool material is placed in the tube, it will become unusable.



7 Insert the stool sampling stick into the tube, then screw the green cap back on. The stool sampling stick is located in the section of the tube that contains liquid, this is a harmless saline solution. The liquid may become a little cloudy. This is normal.



8 Place the tube in the small bag that is provided. This is the white bag with the blue border.



9 Fill in the reply form. Specify the date on which you put the stool sample in the tube. Put the bag and the reply form in the silver-coloured return envelope.

10 Store the return envelope in the refrigerator until you are ready to post it.

11 The return envelope containing the sample and the reply form must be posted **within 24 hours**. You do not need a stamp.

Answers to Frequently Asked Questions

Until what date can I use the tube?
The expiry date is on the tube (year, month).

I have to send the tube on a day the mailbox is emptied. Why is this?
On these days, the period between posting the sample and its arrival at the laboratory is as short as possible. The sooner the laboratory test can be carried out, the better. If you post the sample on Friday evening, Saturday or Sunday, it will not be delivered to the laboratory until Tuesday.

I have lost the tube/the tube is damaged. What should I do now?
If you misplace the tube or if, for example, some of the liquid has accidentally leaked out, you can request a new tube. You can do this by calling the screening organisation's Information Line.

Something went wrong. What should I do now?
Did something go wrong with the tube or do you have a problem collecting the sample? Call the screening organisation's Information Line.

Why do I need to store the tube containing the stool sample in the refrigerator until I can post it?
The laboratory test is better able to detect any blood in your stool if the sample is stored at a low temperature.

Where can I obtain further details?
- A video clip on the website shows the best way to place a stool sample in the tube. This website also has answers to Frequently Asked Questions.
- Call the screening organisation's Information Line. Details of the phone number and the website of the screening organisation in your region are given in the letter or at *www.bevolkingsonderzoekdarmkanker.nl*.

APPENDIX II

Table 1: FOB-Gold Standard for Faecal Immunochemical Tests for Hemoglobin (FITTER) Checklist[21,22]

Topic	Item	Documentation
Specimen collection and handling	Name of specimen collection device and supplier (address)	Name collection device: FOB-Gold Supplier: FOB-Gold, Sentinel Diagnostics, Via Robert Koch, 2-20152 Milan, Italy
	Description of specimen collection device (vial with probe/stick, card, other)	Round tube with collection stick immerged in a preservative solution
	Description of specimens used if an in vivo study (single or pooled feces, artificial matrix with added blood, etc)	Single human faeces sample
	Details of faecal collection method (sampling technique and number of samples)	Ribbed section of the sampling stick is dipped in 4 different parts of the stool
	Who collected the specimens from the samples (patient, technician, etc)	Participant
	Number of faecal specimens used in the study (single, pooled, individual patient feces)	Single sample of individual patient feces
	Mean mass of faeces collected[a]	10 mg
	Volume of buffer into which specimen is taken by probe, applicator stick, or card[a]	1.7 mL
	Time and storage conditions of faecal specimen from "passing" to sampling, including time and temperature (median and range)	Analysis took place at same day of arrival (<24 h) of the FIT in the laboratory and the FIT was kept by ambient air temperature
	Time and storage of collection devices from specimen collection to analysis, including time and temperature (median and range). A concise description of process from collection to analysis is recommended.	Participants were asked to post the faeces samples within 24 h after collection and keep the sample in the refrigerator. The date of sample collection is noted. FIT was transported and analysed by ambient air temperature.
Analysis	Name of analyzer, model, supplier (address), number of systems if more than 1 used.	Bio Majesty JCA-BM6010/C, serial number CA 1401000690069. Supplier: Sysmex Nederland BV Ecustraat 11 4879NP Etten Leur
	Number of times each sample was analysed.	Single or twice. If first analysis resulted in "no results" analysis was repeated.
	Analytical working range[a] and whether samples outside this range were diluted (factor) and re-assayed.	0.4–797.2 ng/mL. Client samples outside this linearity range were not diluted.
	Source of calibrator(s) (supplier with address), number of calibrator(s), how concentrations were assigned[a] and details of calibration process including frequency.	Calibrator supplier: Sentinel Diagnostics, Via Robert Koch, 2-20152 Milan, Italy Calibrator levels: 6 Standard calibration is performed with every reagent and calibrator lot number change.

Table 1: FOB-Gold Standard for Faecal Immunochemical Tests for Hemoglobin (FITTER) Checklist[21,22] (continued)

Topic	Item	Documentation
	Analytical imprecision,[a] ideally with number of samples analysed, concentrations, and mean, SD and CV.	Before the go-live a CLSI EP5A2 protocol was performed on all 3 Sentinel controls (low, mid, high) to verify the imprecision specifications conform tender requirements of the Dutch colorectal cancer screening programme. CLSI EP5A2 results: Sentinel low: 50 ng/mL Lot number control: 30004/A0546 SD with-in (calculated) = 4.04 SD with-in (claim) = 5.00 (10% of 50 ng/mL) (User variance/claim variance) degrees of freedom = 26.13 Critical χ^2 value = 55.76 Claim accepted? Yes SD total (calculated) = 5.12 SD total (claim) = 7.50 (15% of 50 ng/mL) (User variance/claim variance) df = 25.63 Critical χ^2 value = 73.03 Claim accepted? Yes Sentinel mid: 71 ng/mL Lot number control: 30004/A0551 SD with-in (calculated) = 3.69 SD with-in (claim) = 7.10 (10% of 71 ng/mL) (User variance/claim variance) df = 10.82 Critical χ^2 value = 55.76 Claim accepted? Yes SD total (calculated) = 5.49 SD total (claim) = 10.60 (15% of 71 ng/mL) (User variance/claim variance) df = 11.98 Critical χ^2 value = 61.66 Claim accepted? Yes Sentinel high: 312 ng/mL Lot number control: 30004/A0552 SD with-in (calculated) = 4.81 SD with-in (claim) = 31.20 (10% of 312 ng/mL) (User variance/claim variance) df = 0.95 Critical χ^2 value = 55.76 Claim accepted? Yes SD total (calculated) = 7.41 SD total (claim) = 46.80 (15% of 312 ng/mL) (User variance/claim variance) df = 1.08 Critical χ^2 value = 59.30 Claim accepted? Yes

Table 1: FOB-Gold Standard for Faecal Immunochemical Tests for Hemoglobin (FITTER) Checklist[21,22] (continued)

Topic	Item	Documentation
Quality management	Source (address) or description of internal quality control materials, number of controls, assigned target concentrations and ranges, how target concentrations were assigned, rules used for acceptance and rejection of analytical runs.	3 rounds of control before running daily analyses were done and 3 rounds after, conform Sentinel's quality rules. If 2 of 3 controls are within the range, analytical runs are accepted. Apart from the Sentinel controls, a mid-daily run of control conform SKML (SKML CFB, Mercator 1, Toernooiveld 214, NL-6525 EC, Nijmegen, The Netherlands) is performed, every other day a high run or low run: SKML low: 212 ng/mL →5.05% SKML high: 510 ng/mL →3.10% If the control is not right, controls are being repeated, if not right after multiple control rounds, the clinical chemist is consulted.
	Participation in external quality assessment schemes: (name and address of scheme), frequency of challenges, performance attained.	Participation in external quality assessments of SKML (foundation of quality control of medical laboratory diagnostics) following a fixed schedule. Assessment results are monitored by the national functionary FIT
	Accreditation held by the analytical facility (address).	Accreditation by CCKL, Mariaplaats 21-D, 3511 LK Utrecht
	The number, training and expertise of the persons performing the analyses and recording the results	7 trained technician's
Result handling	Mode of collection of data; manual recording or via automatic download to IT system, single or double reading	Results are automatically uploaded to ScreenIT, after authorisation by the laboratory analyst the results are uploaded to the screening IT system ScreenIT.
	Units used, with conversions to µg Hb/g faeces if ng Hb/mL used.	In analyzing and reporting results ng Hb/mL was used. For reporting in publications this is converted to µg Hb/g faeces.
	Cutoff concentration(s) if used and explanation of how assigned locally or by manufacturer	Positive: ≥275 ng Hb/mL (47 µg Hb/g faeces). This was locally assigned by researchers and approved by the Ministry of Health.
	Were the analysts blinded (masked) to the results of the reference investigation and other clinical information?	Yes

APPENDIX III: UNCERTAINTY ANALYSES

Uncertainty Analysis I

Table 1: Multivariable logistic analysis of persons with positive FIT* test result by combined batch-reagent lot and laboratory

Device and reagent lot	Population tested	FIT positives n (%)	OR (95% CI)	*p* value
AAD3	113,283	7,085 (6.3)	REF	<0.001
AAA1	253,600	17,022 (6.7)	0.92 (0.88-0.95)	
AAA2	72,158	4,915 (6.8)	1.05 (0.99-1.12)	
AAA3	1,014	81 (8.0)	1.31 (1.02-1.69)	
AAC1	14,124	915 (6.5)	0.99 (0.91-1.06)	
AAC2	175,844	10,642 (6.1)	0.98 (0.94-1.03)	
AAC3	136,563	8,604 (6.3)	1.02 (0.96-1.07)	
AAC4	1,738	95 (5.5)	0.95 (0.76-1.18)	
AAC5	436	36 (8.3)	1.51 (1.05-2.18)	
AAD4	246,770	15,129 (6.1)	0.98 (0.94-1.02)	
AAD5	168,162	10,486 (6.2)	0.97 (0.92-1.02)	
AAD6	86,017	6,111 (7.1)	1.14 (1.08-1.20)	
AAE56	101,460	6,398 (6.3)	1.00 (0.96-1.04)	
Laboratory				
LabA	442,883	28,651 (6.5)	REF	0.004
LabB	452,776	29,300 (6.5)	1.03 (1.01-1.05)	
LabC	475,510	29,568 (6.2)	1.02 (1.01-1.04)	

Abbreviations: FIT (faecal immunochemical test for haemoglobin).

* FITs were considered positive at a cut-off concentration of 47 µg Hb/g faeces.

**Multivariable OR were corrected for sex, ambient temperature, sample return time and time to expiry of collection devices.

Uncertainty analysis II

Table 2: Multivariable logistic analysis of persons with positive FIT* test result by specimen collection device, reagent lot and laboratory by two age groups

	Age 63		Age 67	
Device	**OR (95% CI)**	***p* value**	**OR (95% CI)**	***p* value**
AAA	REF	0.23	REF	0.004
AAC	0.92 (0.77-1.10)		0.86 (0.76-0.97)	
AAD	0.89 (0.68-1.16)		0.74 (0.60-0.92)	
AAE	0.74 (0.48-1.14)		0.55 (0.39-0.78)	
Reagent lot				
1	REF	0.005	REF	<0.001
2	1.08 (0.95-1.23)		1.19 (1.09-1.29)	
3	1.13 (0.95-1.33)		1.33 (1.17-1.52)	
4	1.21 (0.98-1.50)		1.39 (1.15-1.67)	
5	1.13 (0.86-1.47)		1.46 (1.16-1.83)	
6	1.38 (1.01-1.88)		1.79 (1.39-2.29)	
Laboratory				
LabA	REF	0.69	REF	0.38
LabB	0.98 (0.94-1.03)		1.03 (0.98-1.07)	
LabC	0.99 (0.95-1.04)		1.03 (0.98-1.07)	

Abbreviations: FIT (faecal immunochemical test for haemoglobin).* FITs were considered positive at a cut-off concentration of 47 µg Hb/g faeces.
**Multivariable OR were corrected for sex, age, ambient temperature, sample return time and time to expiry of collection devices.

8

ADHERENCE TO SCREENING IN FIT-BASED COLORECTAL CANCER SCREENING PROGRAMMES IN THE NORTHWEST OF EUROPE

Esther Toes-Zoutendijk, Isabel Portillo, Sarah Hoeck, Isabel de Brabander, Philippe Perrin, Catherine Dubois, Monique E. van Leerdam, Iris Lansdorp-Vogelaar, Marc Bardou

Journal of Medical Screening

ABSTRACT

Background

This study compared adherence to four faecal immunochemical testing (FIT) based screening programmes for colorectal cancer (CRC) in Flanders, France, Basque country and the Netherlands to identify factors to further optimise FIT programmes.

Methods

Background information and data on performance indicators were collected and compared for the four CRC screening programmes.

Results

Invitation method, reminders, funding, FIT cut-off and follow-up after positive FIT differed between the four programmes. In France only an invitation letter is send by mail, while the sample kit needs to be collected at general practitioner (GP). In the other programmes, an invitation letter including the sample kit is send by mail. Participation rates varied substantially with method of invitation, with the highest participation rates in the Netherlands (73.0%) and Basque country (72.4%), followed by Flanders (54.5%) and France (28.6%). Basque country (92.8%) and France (88.4%), the two programmes with most active involvement of GPs in referral for colonoscopy, showed the highest participation rate with colonoscopy.

Conclusions

Large differences in screening participation were observed between programmes in line with the invitation method used. This finding suggests that changes to the design of the programme, such as including the sample kit with the invitation or active involvement of GPs, might increase participation.

INTRODUCTION

Many countries or regions have implemented colorectal cancer (CRC) screening by faecal occult blood testing (FOBT), in particular by means of faecal immunochemical testing (FIT).[1] FOBT as screening method is also recommended by the European Union.[2] The effectiveness of population-based screening programmes is not only driven by the sensitivity of the screening method, but also depends on the availability of resources, healthcare infrastructure and population preferences in each country. Population preferences will especially be reflected in participation rate.

To determine the most optimal screening method for the population, pilot studies were performed in the Basque country, Flanders and the Netherlands before the initiation of the regional or national screening programme. In the Basque Country a pilot study was carried out in 2009 and high participation rate with FIT screening of 64.3% was demonstrated.[3] In Flanders, a pilot study was performed to compare two invitation strategies: FIT directly send by mail or invitation to collect the FIT at the general practitioner (GP). Participation by mail was 52.3% versus 24.6% through the GP.[4] In the Dutch pilot studies different screening methods were compared. These studies showed that FIT screening resulted in the highest CRC detection per invitee compared to other screening methods like guaiac FOBT (gFOBT), colonoscopy and sigmoidoscopy screening.[5-7]

After choosing the best screening method, the organisational structure of the programme is crucial for optimal screening performance. Many different aspects how to organise the programme have been studied like pre-invitation letter, reminders, and FIT mailing.[4,8,9] However, almost all of these studies have been carried out in trials and not in real-life settings. In running programmes many more aspects are involved, for example organisation of healthcare systems and healthcare insurance. Besides, these organisational aspects have never been compared across programmes, but only in one specific group of individuals. It is unknown of all these different organisational aspects will work out the same for individuals residing in different countries. A national population-based CRC screening programme was initiated in 2002 in France using gFOBT, which was changed to FIT in April 2015. FIT screening was introduced in 2009 in the Basque country (Spain), in 2013 in Flanders (Belgium), and in 2014 in the Netherlands. As these programmes are geographically close and connected, all situated in Europe, and have recently been implemented similar outcomes with respect to CRC screening may have been expected. This study was able to evaluate similarities and differences between the organised population-based CRC screening programmes using FIT in France, Basque country (Spain), Flanders (Belgium), and the Netherlands and assesses how this may impact adherence to FIT-based programmes.

METHODS

Organisational structure of CRC screening programmes

Information was collected on year of initiation, target population, eligible population, screening interval, methods of invitation to FIT screening and to colonoscopy following a positive FIT, funding and executive organisation of the screening programmes. The target population was defined for each population-based CRC screening programme according to programme specific policies. The eligible population is the target population excluding those that are not eligible for screening based on exclusion criteria. Eligible population were all individuals that should have been invited in 2016. This number can deviate from the total target population, because of biennial screening or phased implementation of the national screening programme.

Performance indicators of CRC screening programmes

Data on performance indicators were extracted from each of the national or regional screening databases. Data from France were extracted from the database of French Public Health Agency (Santé Publique France) and Organized screening structure of the Big East region and the Pyrénées. Data from Flanders were extracted from the screening database, the Belgian Cancer Registry and reimbursement data from the Health insurance companies. All data from the Basque country were extracted from programme database (PCCR) which is linked with medical records, population and hospital cancer registries. All data from the Netherlands were extracted from the national database for screening programmes (ScreenIT). In France, data on the invitees starting in April 2015 to December 2016 were collected until June 2017, in the Basque country data on the invitees of 2016 were collected until December 2017, in Flanders and the Netherlands data on the invitees of 2016 were collected until 30 June 2017.

Data were collected on main performance indicators: participation rate, positivity rate of the FIT, participation rate to colonoscopy following positive FIT, detection rate of CRC or advanced neoplasia (AN) per participant, and diagnostic yield. Definitions of the indicators are in accordance with recommended definitions for performance indicators by the European Union CRC screening guidelines.[11]

1. *Participation rate* was calculated as the number of persons sending back the FIT sample divided by the number of persons receiving an invitation letter. For Flanders and France persons were only considered as participant if they returned the FIT sample within 12 months after the invitation. In the Basque Country persons were only considered participant if they returned an assessable FIT sample within six months after the invitation. In the Netherlands individuals were considered participant until the date of the invitation of subsequent screening round.

2. *Positivity rate* was calculated as the number of persons with a FIT result at or above the cut-off level divided by the number of persons with an assessable stool sample.
3. *Participation rate colonoscopy* was calculated as the number of persons undergoing a colonoscopy divided by the number of persons with a positive FIT result.
4. *Detection rate* was defined as the number of persons with AN detected during colonoscopy per participant. AN was considered as relevant abnormality within a CRC screening programme. AN was defined as CRC or any adenoma with histology showing >=25% villous component or high-grade dysplasia or adenoma with size >= 10 mm. In Flanders, only adenoma with any villous component and/or high grade dysplasia was counted as advanced adenoma, because no data were available on adenoma size or the amount of villous components. In the Basque country, in addition to histology, dysplasia and size, having ≥3 adenomas was also considered as advanced adenoma.
5. *Diagnostic yield of the programme* was defined as the number of persons with AN detected during colonoscopy divided by all individuals that received an invitation. In Flanders, data of colonoscopy yield is not linked to the date of invitees of the programme. The denominator can contain individuals invited in previous year.

Analysis

First, organisational structure of the four programmes were compared using thematic analysis to identify similarities or differences. Second, outcomes of the performance indicators for each of the four programmes were compared. To rule out that the observed difference is related to cultural differences between populations rather than organisational differences, the programme of Basque country in Spain was compared with the Basque country in France. These are two regions that are very close with respect to geographical location and cultural background. The different subgroups were compared using chi-squared test and *p* value <0.05 was considered statistically significant. Chi-square test was performed using R version 3.5.0.

RESULTS

Organisation of the programmes

The age range of the target population differed, with France and the Basque country having the lowest starting age of 50 years and the Basque country having the lowest stopping age of 69 years (Table 1). All four countries used a two years screening interval. Exclusion criteria prior to invitation differed between the four programmes, with very limited exclusion criteria in the Netherlands compared to the other three programmes: persons were only excluded based on a positive FIT in previous screening round (Table 1). France, Flanders and Basque country all excluded individuals with history of CRC, proctocolectomy, and recently

Table 1: Background information on screening programme

	France	Flanders (Belgium)	Netherlands	Basque country (Spain)
		1. General		
Year of introduction	Started in 2009 and switched nationally to the FIT in April 2015.	October 2013 Phased implementation by age groups. In 2015 the programme was fully implemented.	2014 Phased implementation by age group. In 2019 the programme will be fully implemented.	2009 Phased implementation by provinces. In 2014 the programme was fully implemented.
Piloting	Yes	Yes	Yes	No
Public Awareness Campaigns	Yes. Through dedicated websites: National Cancer Institute (INCa), Health Insurance, and some local initiatives.	Yes. Since 2015 short announcements on the Flemish television and since 2017 advertising in public transport. Both only during the month March (international CRC month).	Yes. Information on the website of the national institute of public health and the environment. Only at the start in 2014 there was extra media attention through television and radio. No advertising.	Yes. National yearly campaign by patient association and main results are presented in television, leaflets, radio, and social network.
Population-based programme	National	Regional	National	Regional
Age group	50-74 years	56-74 years Per July 1 2017 it was extended to 55-74 years and per July 2018 it was extended to 53-74 years	55-75 years	50-69 years
Screening interval	2 years	2 years	2 years	2 years
Cut-off level (Hb/g faeces)	30 µg	15 µg	47 µg	20 µg
Brand name of the test	OC-Sensor, Eiken, Japan	OC-Sensor, Eiken, Japan	FOB-Gold, Sentinel, Italy	OC-Sensor, Eiken, Japan
		2. Methods of invitation		
Pre-invitation letter	No	No	Yes, 3 weeks in advance	Yes, 4 weeks in advance
Invitation by mail including FIT	No, test to be collected at GP's office	Yes	Yes	Yes
Reminder letter	12 weeks and 24 weeks	8 weeks	6 weeks	4 weeks
Exclusion of individuals before invitation	Yes	Yes	No	Yes

Table 1: Background information on screening programme (continued)

	France	Flanders (Belgium)	Netherlands	Basque country (Spain)
Exclusion criteria mentioned in the invitation letter	No.	Past or current treatment of colorectal cancer, occult blood in stool, and unexplained and persistent change in the bowel movement patterns, colonoscopy in past ten years, stool test in past 2 years, higher risk at CRC, having one or more first relatives who have/had CRC.	Past or current treatment of colorectal cancer, occult blood in stool, and unexplained and persistent change in the bowel movement patterns.	Colonoscopy performed ≤ 5 years.
		3. Methods of invitation to diagnostic colonoscopy		
Invitation to follow-up colonoscopy	By letter. Results are sent to the participant, the general practitioner and the structure in charge of organised CRC screening	By letter. Results are sent to the participant and the general practitioner.	By letter, including appointment for colonoscopy intake. Results are sent to the participant and if possible the general practitioner.	By letter. Results are sent to the participant. In that letter the participant is recommended to visit GP for colonoscopy referral.
Reminder following initial invitation colonoscopy	No	No, from 2019 persons receive a new invitation for colonoscopy intake 2 years after their positive FIT. When not following this second advice, 2 years later (4 years after the positive FIT) they will receive a new FIT.	Yes, persons receive a no show letter. In case of no response, persons received a new invitation for colonoscopy intake after 2 years after positive FIT (until 2017). From 2018 onwards persons will receive a new FIT after two years.	Yes, after 30 days the participant is reminded to make an appointment with the general practitioner for colonoscopy referral. If participants reject to have a colonoscopy follow-up they will receive a new FIT after 2 years.

Table 1: Background information on screening programme (continued)

	France	Flanders (Belgium)	Netherlands	Basque country (Spain)
		4. Organisation		
Funding of screen test	Free of charge, but not the GP's encounter to collect the kit that is covered at 70% by national public health insurance. The remaining 30% are being covered by complementary insurance if the participant has one.	Free of charge	Free of charge	Free of charge
Funding of colonoscopy	Standard healthcare insurance covers 70%. The remaining 30% is covered by complementary insurance if any.	Standard healthcare insurance and partially payment of personal funds (out of pocket costs).	Standard healthcare insurance. Note: Not in all situations colonoscopy costs are fully covered by the healthcare insurance. Up to 350 - 850 euro are paid of personal funds (out of pocket costs, only once a year a deductible excess of all healthcare costs).	Free of charge
Responsible organisation of the screening programme	INCa, National cancer institute, with dedicated screening structure in each of the departments. The latest are in charge of invitations.	Centre for Cancer Detection (CvKO). In Flanders CvKO consists of five regional screening programmes, but the invitation letter, reminder letter, and FIT results are organised by one central organisation. The five regional screenings programmes offer free telephone advice and are responsible for administration.	The national institute of public health and the environment is the responsible organisation of the programme. Five regional screening organisations are responsible for the execution of the programme. Those five regional screening organisations are brought together in one national cooperation.	The Basque Health Service is responsible for the execution of the programme, according with authorities planning, organising and monitoring the quality of the process and results. The final responsibility of the programme is the Regional Ministry of Health in the Basque Country.

performed colonoscopy before invitation. France and Flanders also excluded individuals with a recently performed FIT. Additionally, the Basque country excluded individuals with severe or terminal illness.

Methods of invitation differed between the four programmes. The eligible population in France received an invitation letter to collect the FIT sample kit at the GP. In Flanders individuals received an invitation including the FIT sample kit. In the Basque country and the Netherlands a pre-invitation letter was send prior to invitation followed by an invitation letter including the FIT sample kit. All four screening programmes used a reminder letter, but all at different time points ranging from 30 days (Basque country) until 6 months (France). France sent two reminder letters. All four programmes used another cut-off for a positive FIT for referral to colonoscopy: Flanders used the lowest cut-off of 15 µg Hb/g faeces, followed by the Basque country with 20 µg Hb/g faeces and France with 30 µg Hb/g faeces. The highest cut-off was used in the Netherlands with 47 µg Hb/g faeces (Table 1).

Performance indicators

A total of 18.9 million individuals were invited to participate in FIT screening among the four CRC screening programmes. Highest participation rate was observed in the Netherlands (73.0%), followed by the Spanish Basque country (72.4%), Flanders (54.5%) and France (28.6%, p <0.001). As a consequence of the different FIT cut-offs used, positivity rate differed between the four programmes, from 4.7% in France to 6.7% in Flanders (p <0.001). Highest participation rate for the colonoscopy following a positive FIT result was observed in the Basque country (92.8%), France (88.4%), Flanders (81.9%) and the Netherlands (82.8%) (p <0.001). Detection rate for AN per participant was highest in the Netherlands (2.3%) and lowest in Flanders (1.0%). Diagnostic yield for AN per invitee was highest in the Netherlands (1.6%) and lowest in Flanders (0.6%, p <0.001).

French versus Spanish Basque country

Despite cultural similarities, differences in screening performance indicators were observed between the French and the Spanish parts of the Basque country (Table 3). The participation rate in the Spanish part, with 72.4% was 2.5 times as high as the French part of the Basque country, with 24.6% (p <0.001; Table 3). Participation rate to colonoscopy was of the same magnitude in both regions: 92.8% in the Spanish part and 87.4% in the French part (p 0.37).

DISCUSSION

Large differences in screening participation were observed between programmes in line with invitation method used, such as a pre-invitation letter and including the FIT sample kit with the invitation. The high participation to colonoscopy in France might indicate that well

Table 2: Performance indicators for France, Flanders, the Netherlands and Basque country

	France	Flanders	Netherlands	Basque country	*p* value
Calendar year	2015-2016	2016	2016	2016	
Age (year)	50-74	56-74	59-76	50-69	
Target population	19,043,771	1,447,434†	Unknown	273,084	
Eligible population	16,701,387	830,665	1,543,223	239,601	
Invited	16,701,387 100%	571,034 68.7%†	1,457,976 94.5%	229,380 87.7%	
Number of participants	4,779,845	311,453	1,063,651	166,110	<0.001
Participation rate FIT	28.6%	54.5%‡	73.0%	72.4%	
Men	*27.8%*	*53.1%*	*71.1%*	*70.0%*	
Women	*30.8%*	*56.0%*	*74.8%*	*74.6%*	
Screen round	Any round	First and second	First and second	First to Fourth	
Cut-off level (Hb/g faeces)	30 µg	15 µg	47 µg	20 µg	
Positivity rate	4.7%	6.7%	5.4%	5.2%	<0.001
Participation rate colonoscopy	88.4%*	81.9%	82.8%	92.8%	<0.001
Detection rate					
AN	1.5%*	1.0%	2.3%	1.9%	<0.001
CRC	0.31%*	0.28%	0.35%	0.20%	<0.001
Diagnostic yield programme					
AN	0.4%*	0.6%	1.6%	1.4%	<0.001
CRC	0.09%*	0.15%	0.25%	0.15%	<0.001

† Eligible population in Flanders is the total amount of 56-74 years old for two year minus those excluded for invitation. Eligible population for 2016 only could not be provided.

‡ Coverage by examination, also including opportunistic screening by FIT or colonoscopy, resulted in 65.5% of the target population to be screened.

* In France the participation rate of colonoscopy and number of colorectal cancers and advanced neoplasia was based on data from April 2015 until December 2015.

Abbreviations: AN (Advanced Neoplasia); N.A. (Not available).

Advanced neoplasia was defined as CRC or any adenoma with histology showing >=25% villous component or high-grade dysplasia or adenoma with size >= 10 mm. In Basque country also ≥3 adenomas were considered AN. In Flanders, adenoma with a villous component and/or high grade dysplasia was counted as advanced adenoma. There were no data available on the size or the amount of villous components in an adenoma. Detection rate, invitees with CRC or AA per participant. Diagnostic yield, individuals with CRC or AN per invitees.

informed and motivated people that collect the FIT sample kit at the GP, are more likely to undergo a colonoscopy.

For the large difference in FIT participation we have several explanations. First, sending the FIT home is more effective than collecting it at the GP. Almost all studies were irrevocably showing a huge increase in participation when including the FIT sample kit with the invitation.[8,12-14] However, one Italian study showed only a modest increase in participation, but this study was performed in previously screened individuals (used to other screening

Table 3: Outcomes performance indicators Basque region

	Basque country in France	Basque country in Spain	*p* value
Year	2016	2016	
Age	50-74	50-69	
Invited	45,923	229,380	
Number of participants	11,293	166,110	
Participation rate FIT	24.6%	72.4%	<0.001
Cut-off level (Hb/g faeces)	30 µg	20 µg	
Positivity rate	4.6%	5.2%	0.07
Participation rate follow-up colonoscopy	87.4%	92.8%	0.37
Detection rate			
AN	1.4%	1.9%	<0.001
CRC	0.27%	0.20%	
Diagnostic yield			
AN	0.4%	1.4%	<0.001
CRC	0.07%	0.15%	

Abbreviations: AN (Advanced Neoplasia)

Advanced neoplasia was defined as CRC or any adenoma with histology showing >=25% villous component or high-grade dysplasia or adenoma with size >= 10 mm. In Basque country also ≥3 adenomas were considered AN. Detection rate, invitees with CRC or AN per participant. Diagnostic yield, individuals with CRC or AN per invitees.

strategy).[16] One French study showed low uptake rates with direct mailing of the FOBT.[16] This inconsistency may be due to the test modality, gFOBT instead of FIT, resulting in lower participation rates.[17] Second explanation for a higher FIT participation may be the advanced notification letter as illustrated by the higher participation rate in the Basque country and the Netherlands. However, this will only explain a small proportion of the total difference, as studies have shown that sending a pre-invitation letter results in a three percentage point increase.[9,10] Only one study from Australia showed a higher increase, nine percentage point.[18] Both direct mailing as well as the pre-invitation letter are in line with a recent systematic review.[19] However, one large difference is noteworthy. The review reported that GP involvement improved participation. We showed the opposite in this study; a country with no involvement of GPs like the Netherlands, participation rates were very high, while in a country with active involvement of the GPs like in France, participation rates were substantially lower. We hypotheses that GP endorsement can have a positive impact on participation, as long this requires no effort of the participant. This is in line with findings of the CRC screening programme in England, showing an increase in participation if the invitation letter was added with a GP endorsement banner. Our analysis of the two Basque regions in France and Spain showed that very similar cultures can have very different rates in screening participation, and that culture may not be the driving factor of performance

differences between programmes. However, we cannot rule out cultural differences completely. We know from literature that cultural difference in screening attitude is also observed in the participation rates of other cancer screening programmes, for example participation to breast cancer screening. In 2016, this was also lower in Flanders (51.9%) than in the Netherlands (77.6%) and the Basque country (80.1%), with France having the lowest participation rate (50.7%).[21-23] Remarkably, the participation rate for breast cancer screening in France is similar to Flanders, while there is a much larger difference in participation rate for CRC screening. Gender cannot explain this difference, as both men and women showing a similar pattern in participation. Thus, this again reflects the negative impact of using a different invitation method in France for CRC screening.

Participation rate to follow-up colonoscopy was considerable high in all four screening programmes. However, the rate was below the recommend level of 85% in the Netherlands and Flanders. We hypothesize that higher participation to follow-up colonoscopy can be the result of the active involvement of GPs during the screening process. In France and the Basque country GPs play an active role in 1) defining the eligible population by excluding those with severe comorbidity from invitation, 2) selecting the population eligible for FIT screening at pick-up of the screening test or 3) following individual up after negative FIT. Consequently, those participating in FIT screening are all healthy enough to undergo follow-up colonoscopy. Other way around it also explains the lower participation to colonoscopy in the Netherlands, as there is no exclusion of individuals based on co-morbidities or medical history. Additionally, in France probably only the most motivated individuals collect the FIT sample kit at their GP practice and they may be more motivated to go for colonoscopy in case of a positive FIT. Only involving GPs for referral to colonoscopy, without involvement in selecting those eligible for FIT screening, will be less effective.[8] Reimbursement differences of the colonoscopy do not seem to explain participation differences. Although in the Basque country the colonoscopy is free of charge, the participation in France was only slightly lower, while French individuals may have significant expenses.

Positivity rate differed for all the four programmes. This is due to three important reasons: cut-off of the FIT, target age group and screening round (first or subsequent round).[24] The same explanations hold for the difference in detection rates and diagnostic yield of the programme. We could not restrict our analysis for the same age ranges, as the Netherlands is still in the implementation phase and not all age groups of the target population have been invited yet. Therefore, the outcomes of the positivity rate and detection rates should be addressed as exploratory, and further research is needed to explain the differences between these rates.

Our study has three strengths. It is the first that gives detailed information on organisational structure of four programmes provided by representatives of each country. These details are in general unknown, as key elements of CRC screening programmes are only described in its own country language: Flemish, French, Spanish/Basque, and Dutch. These details can

be used by other countries/regions considering CRC screening and are valuable for policy makers. Also, our study showed very recent outcomes of four large population-based programmes, all using the same test modality (FIT). Lastly, our study compared screening programmes of neighboring countries with cultural similarities and differences, and can thus address the impact of cultural and organisational aspects in the uptake of CRC screening.

The study has also some limitations. First, comparing quality indicators was challenging due to different definitions and differences in cut-off and number of screening rounds. Unfortunately, we could not restrict the comparison to first screen round data only as not all programmes had such detailed information. Second, data collection may be of a concern, for example France does not have a central data collection of quality indicators and diagnostic yield.

The findings of the study suggest that the organisational structure impacts the participation rate to FIT and follow-up colonoscopy, like sending out the FIT, pre-invitation letter, involvement of the GP in the whole screening process. These results can be used to optimise each of the four screening programmes or can be used as an example for other organised FIT-based CRC screening programmes. Possibilities for optimisation can be diverse for every programme as health care systems, funding of the colonoscopy and available resources differ. Interventions for optimisation will cost money and these results can therefore be used to explore the additional benefit and additional costs for each of the programmes. France already started optimising their screening programme, but maybe not in the good way. Indeed it has been decided to mail the FIT with the first reminder but only to those who had already been participating in previous round, whereas the study by Giorgi-Rossi and colleagues suggests that they may not be the best target.[15]

Although sending the FIT by mail and actively approaching FIT positives for the colonoscopy seems to be most effective, this can be considered as infringement of free will.[25] High participation should not be the goal of screening programmes, but the level of informed choice. However, there is no indication that high participation in the Netherlands for example, results in a lower level of informed choice.[26,27] Besides these ethical considerations, there is also a remaining difference in participation that cannot be explained by the organisational structure and is difficult to unravel. It seems to be a difference in attitude towards screening in general between the different regions or countries. It is unclear how this arises and can be solved.

In conclusion, this study shows that including the FIT with the invitation results in higher FIT participations rates. Actively involvement of the GP will result in higher participation rates to colonoscopy follow-up, but only if no effort of participants is required. Adjustments to the organisational structure of a screening programme may result in more screening benefit.

REFERENCES

1. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49.
2. Luxembourg: Council of the European Union. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri-OJ:L:2003:327:0034:0038:EN:PDF. [Cited September 21 2017].
3. Portillo I, Idigoras I, Ojembarrena E, Arana-Arri E, Zubero MB, Pijoan JI, et al. Principales resultados del programa de cribado de cancer colorrectal en el Pais Vasco. Gac Sanit. 2013;27(4):358-61.
4. Van Roosbroeck S, Hoeck S, Van Hal G. Population-based screening for colorectal cancer using an immunochemical faecal occult blood test: a comparison of two invitation strategies. Cancer Epidemiol. 2012;36(5):e317-24.
5. 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.
6. 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.
7. 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.
8. Rat C, Latour C, Rousseau R, et al. Interventions to increase uptake of faecal tests for colorectal cancer screening: a systematic review. Eur J Cancer Prev. 2018; 27(3):227-236.
9. Senore C, Ederle A, DePretis G, et al. Invitation strategies for colorectal cancer screening programmes: The impact of an advance notification letter. Prev Med. 2015;73:106-11.
10. 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.
11. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on Cander. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation of screening outcomes. Endoscopy. 2012;44 Suppl 3:SE49-64.
12. Piette C, Durand G, Bretagne JF, et al. Additional mailing phase for FIT after a medical offer phase: The best way to improve compliance with colorectal cancer screening in France. Dig Liver Dis. 2017;49(3):308-11.
13. Ponti A, Anttila A, Ronco G, et al. Cancer Screening in the European Union. Report on the implementation of Council Recommendation on Cancer Screening. Brussels: European Commission; 2017. Available from : https://ec.europa.eu/health/sites/health/files/major_chronic_diseases/docs/2017_cancerscreening_2ndreportimplementation_en.pdf. [Cited October 24 2018].
14. Vanaclocha-Espi M, Ibanez J, Molina-Barcelo A, et al. Factors influencing participation in colorectal cancer screening programs in Spain. Prev Med. 2017;105:190-6.
15. Giorgi Rossi P, Grazzini G, Anti M, et al. Direct mailing of faecal occult blood tests for colorectal cancer screening: a randomized population study from Central Italy. J Med Screen 2011;18: 121-7.
16. Denis B, Gendre I, Perrin P. Participation in four rounds of a French colorectal cancer screening programme with guaiac faecal occult blood test: a population-based open

cohort study. J Med Screen. 2015;22(2):76-82.

17. 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. 2017;66(9):1631-1644.
18. Cole SR, Smith A, Wilson C, et al. An advance notification letter increases participation in colorectal cancer screening. J Med Screen 2007;14: 73-5.
19. Duffy SW, Myles JP, Maroni R, et al. Rapid review of evaluation of interventions to improve participation in cancer screening services. J Med Screen 2017;24: 127-45.
20. Wardle J, von Wagner C, Kralj-Hans I, et al. Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the English NHS Bowel Cancer Screening Programme (ASCEND): four cluster-randomised controlled trials. Lancet 2016;387: 751-9.
21. The Spanish screening cancer network. Available in: www.cribado de cancer.es [Accessed January 13, 2018].
22. Monitoring report of the Flemish Colorectal Cancer Screening Programme 2017. Available from: https://dikkedarmkanker.bevolkingsonderzoek.be/sites/default/files/atoms/files/Jaarrapport2017_DEF_0.pdf [Cited January 13 2018}.
23. Monitoring report of the French screening programme 2016-2017. Available from: https://bit.ly/2HIc0BP. [Cited April 13 2018).
24. Arana-Arri E, Idigoras I, Uranga B, al. Population-based colorectal cancer screening programmes using a faecal immunochemical test: should faecal haemoglobin cut-offs differ by age and sex? BMC Cancer. 2017;17(1):577.
25. Hofmann B. Ethical issues with colorectal cancer screening-a systematic review. J Eval Clin Pract. 2017;23(3):631-41.
26. de Haan MC, de Wijkerslooth TR, Stoop E, et al. Informed decision-making in colorectal cancer screening using colonoscopy or CT-colonography. Patient Educ Couns. 2013;91(3):318-25.
27. van Dam L, Korfage IJ, Kuipers EJ, et al. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? Eur J Cancer. 2013;49(10):2321-30.

9

GENERAL DISCUSSION

The general aim of this thesis was to evaluate the performance of the national colorectal cancer (CRC) screening programme with biennial faecal immunochemical testing (FIT) in the Netherlands during the implementation phase. In this chapter, I first will discuss the most important findings of this thesis. Secondly, I will discuss the methodological considerations of the studies described in this thesis. Finally, I will elaborate on future perspectives and draw final conclusions.

9.1 MOST IMPORTANT FINDINGS OF THE DUTCH COLORECTAL CANCER SCREENING PROGRAMME DURING THE IMPLEMENTATION PHASE

From 2014 until 2017, 5.3 million individuals were invited to participate. A total of 3.9 million individuals returned their FIT. The FIT was positive for 223,043 individuals. A total of 180,398 individuals underwent a colonoscopy. During colonoscopy, 14,084 CRCs and 76,022 AA were diagnosed. Based on these promising results during the implementation phase, it is expected that on the long-term CRC-related mortality will decrease, which is observable by 2029. Below the most important performance indicators of the Dutch CRC screening programme will be discussed separately, combining results of different chapters of this thesis.

Participation rate FIT

In **Chapter 2** we observed a very high participation rate (71.3%) in the first year of the national CRC screening programme. In the study described in **Chapter 4** we evaluated participation in the second screening round and estimated a consistent participation of 93% among the individuals that previously participated. The high FIT participation rate is still one of the highest across the world, as was also confirmed in **Chapter 8**.[1] Such high participation rate to primary screening test offered by the Dutch government is also observed in the two other large cancer screening programmes; 78.8% for breast cancer and 64.6% for cervical cancer.[2,3] High participation rate is a relevant finding, as high participation will eventually result in high CRC detection rates per invitee.[4-6] In the pilot study a stable FIT participation between 60-62% was observed over four screening rounds.[7] If the results of the pilot study reflect what will take place in the national screening programme, the participation rate will remain high in coming years.

Despite the high participation in screening in the Netherlands, still almost 30% of the population does not participate in the CRC screening programme. The question is how much effort should be put in reaching out to nonparticipants. Striving for 100% uptake of the primary screening test should in my opinion not be the goal. It is of great importance that all individuals participating in screening do make an informed choice. The goal should be to increase the number of individuals making an informed choice to participate in screening. We know from previous studies among the Dutch population, that reasons for

nonparticipation are often related to lack of knowledge about CRC.[8,9] Therefore, reaching out to nonparticipants is beneficial and should be undertaken. Up to now, it is unknown how to approach this in the Netherlands.

FIT positivity

In chapter 2, we described that weekly monitoring revealed that FIT positivity rate in the Dutch CRC screening programme was higher than anticipated, with an age-adjusted positivity rate of 10.7% at the cut-off of 15 µg Hb/g faeces. Consequently, a higher number of individuals were referred for colonoscopy. This higher positivity rate together with the lower than anticipated PPV for CRC and AA, resulted in the urgency to adjust the CRC screening programme. The outcomes of the programme were evaluated with a decision analysis tool, aiming to identify the best option to optimise the screening programme. Three options were evaluated: increase cut-off, postpone screening in selected age groups and forego screening in older age groups. This analysis showed that increasing the cut-off level not only resulted in lowest decrease in CRC deaths prevented, but also resulted in a balance between harms and benefits of screening in accordance with that aimed for at start of the programme.[10] The age-adjusted positivity rate of 6.7% at the higher FIT cut-off was now in line with the expected positivity rate of 6.4%.

The higher than anticipated positivity rate and subsequent adjustment of the FIT cut-off was widely debated. How could it possible that after extensive preparations, the national CRC screening programme differed so considerably from expectations? The debate was mostly on similarity of different FIT brands. The FIT brand selected through public tender was FOB-Gold (Sentinel, Italy), which differed from the brand OC Sensor (Eiken, Japan) that was used in the pilot studies. Because equal performance of the two FIT brands was uncertain, accuracy of the two FIT brands were compared.[11] Main finding of this study was that faecal Hb concentrations and FIT positivity differed, but similar detection of CRC or AA at a pre-set positivity rate was observed. It is of note that, although FIT positivity rate differed, OC-Sensor had higher positivity rates at lower FIT cut-offs. Therefore, using FOB-Gold instead of OC-Sensor is not the explanation for the higher positivity rate at the start of the programme. Another more likely explanation is that the manufacturers of the stool tests have improved the test itself. Several programmes had shown that the FIT had a worse performance at higher ambient temperature.[12,13] In a response to this unfavourable outcome, we assumed that the buffer of the test has been improved resulting in better preservation of faecal Hb. Other recent studies also using new generation FITs showed no impact on clinical outcome at higher ambient temperatures and delayed sample return time, indicating improved FIT performance.[14,15] One of these studies proved in laboratory a better Hb stability using FITs with improved buffer.[15]

Participation to follow-up colonoscopy

The high FIT participation rate in the current Dutch programme did not apply to participation to follow-up colonoscopy. The observed participation to the colonoscopy follow-up after a positive FIT of 77.8% in the Netherlands is below the recommended acceptable level of 85%, which is a major concern.[16] In **Chapter 8** we observed a higher participation rate to follow-up colonoscopy in surrounding countries. One explanation for the low participation may be that not all colonoscopy results were integrated in ScreenIT (6-8% of all FIT positives), because some individuals may have had a colonoscopy in centres outside the screening programme. Still, effort should be undertaken to increase this low participation rate, because individuals without appropriate follow-up after a positive FIT were seven times more likely to die from CRC than individuals with appropriate follow-up.[17] Higher participation rates could possibly be reached by active involvement of the general practitioner (GP), as described in chapter 8. Especially the involvement of GPs is subject of continuous debate in the Netherlands. At the start, GPs received the result of individuals with a positive FIT if participants entered GP details in the reply form. Since 2017, a reply form is no longer included and consequently GP contact information is often not provided anymore. This situation makes involvement of GPs to increase colonoscopy follow-up difficult. Currently, options to automatically obtain individuals GPs details from existing databases are explored. If legally and technically possible, the FIT result can automatically be sent to the GP which may have a positive impact on the number of individuals with a complete follow-up after positive FIT.

Involvement of the GP would only lead to an increase in colonoscopy participation, if individual's motives for nonparticipation are unjustified. In 2017 a qualitative study using interviews among Dutch invitees was carried out. This qualitative study showed a wide variety of motives for nonparticipation: low risk perception for CRC, alternative explanation for blood loss, not realising consequence of positive FIT, resentment against colonoscopy, aversions to organisational structure, or unwilling to visit a hospital (Bertels et al. *submitted*). The authors concluded that based on these outcomes increasing individual's risk-perception for CRC might be the most effective to increase colonoscopy participation rate, but needs to be further studied. Potentially, GPs can play an important role in explaining risks of CRC to their own patients. Besides individuals' motive, co-morbidities may often be the reason for nonparticipation. However, individuals are not excluded prior to invitation based on their medical history.

Colorectal cancer and advanced adenoma detection

As discussed above, in Chapter 2 we described that the screening programme was optimised by increasing the FIT cut-off. This was predominantly decided because of a lower than anticipated PPV. As described in **Chapter 1**, the Health Council preferred a FIT cut-off of 15 µg Hb/g faeces over 10 µg Hb/g faeces, aiming for a more optimal balance between true and false positives. The PPV as observed in the first half year of 2014 of 42.1% was below

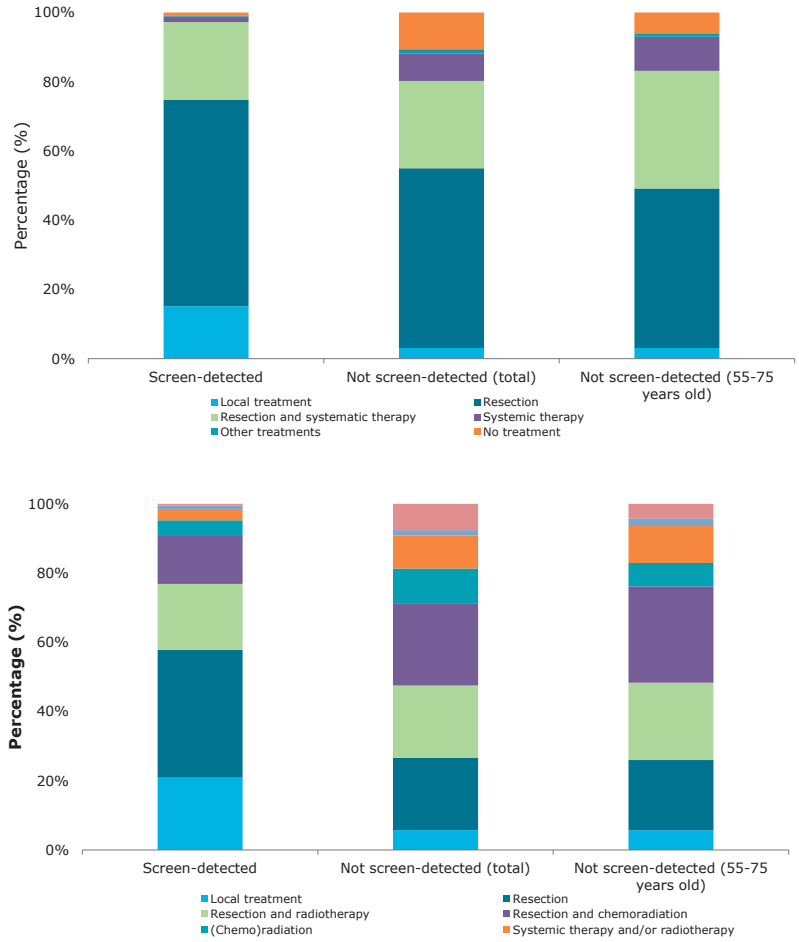
this desired PPV of 51.6%. As a result of the increased FIT cut-off, the PPV for CRC and AA increased to a more desirable level of 49.1%. However, the detection rate decreased for CRC (5.8‰ to 4.4‰) and AA (30.8‰ to 20.6‰), indicating that CRCs will be missed in the first screening round. We hypothesised in Chapter 2 that missed CRCs or AAs in the first screening round, may be detected in the second screening round. In **Chapter 4** we therefore evaluated the impact of the increased FIT cut-off on outcomes of the second screening round. We concluded that using a higher FIT cut-off has limited impact on the CRC and AA detection because a substantial part of the missed lesions will be detected in subsequent screening round. However, after two screening rounds the cumulative yield of CRC and AA is still lower for those tested with the higher cut-off in the first year compared to those tested with the lower cut-off in the first year. It is expected that this difference will become insignificant after more screening rounds. The limited impact of the increased cut-off is also confirmed in **Chapter 5** indicating a small difference in sensitivity between the two FIT cut-offs. Based on both the outcomes of the second screening round and number of interval CRCs, we can cautiously conclude that indeed programme performance has been optimised by increasing the FIT cut-off. Currently, there is no urgent need to change the FIT cut-off. However, it is important to obtain additional information on stage distribution in subsequent rounds, to ensure that CRCs detected in the second round still have a favourable stage distribution at both FIT cut-offs.

In Chapter 4 we also showed that individuals with a faecal Hb concentration between 15-47 µg Hb/g faeces (below FIT cut-off) were 23.2 more likely to have a CRC or AA detected in the consecutive screening round than individuals with no detectable faecal Hb. This makes the previous faecal Hb concentration an important risk factor, which could potentially be used for personalised screening as will be discussed in more detail in the next section.

Stage distribution

In **Chapter 3,** stage distribution of CRCs were compared, showing a more favourable stage distribution (stage I and II) in screen-detected CRCs (66.7%) than in symptom-detected CRCs (39.8%). These findings are in line with expectations, aiming for early detection of cancers thereby improving survival. The results are a promising sign that CRC screening may decrease CRC-related morbidity and eventually mortality rates. Nevertheless, stage distribution of screen-detected CRCs in subsequent screening rounds should be monitored closely to make sure that majority of screen-detected CRCs will still be detected in an early stage. The more favourable stage distribution had also a direct impact on the treatment. Since the introduction of screening in the Netherlands a shift in treatment options was observed for individuals with CRCs detected through screening (Figure 1).[18] Screen-detected CRC patients on average received less invasive and curative treatment compared to symptom-detected CRC patients. This result is satisfying and is an indication that CRC-related morbidity will decrease in coming years.

Figure 1: Treatment options of screen-detected versus non screen-detected (a) colon cancers or (b) rectal cancers in 2015 in the Netherlands.



Location

In Chapter 3 we showed that screen-detected CRCs were more often located in the left hemi colon compared to those CRCs detected without screening. Results of previous studies showed conflicting results.[19,20] Results from Chapter 5 however confirm our hypothesis of a lower FIT sensitivity for right-sided cancers, as many more interval cancers were located right-sided. First explanation might be that FIT is less sensitive for right-sided cancers due to degradation of Hb during colon transit.[21] Another explanation might be that FIT is less sensitive for sessile serrated lesions and this type of polyps are more often detected in the right colon. If these precancerous stages are missed with FIT screening, an increase in proportion of right-sided cancers will appear in the long-term.[22,23] However, this latter

hypothesis can only explain a part of the finding, as approximately 20-30% of the CRCs are thought to derive of the serrated neoplasia pathway. Probably it is a combination of hypotheses, longer transit and lower sensitivity for serrated lesions which makes FIT less suitable for right-sided lesions.

Interval cancers

In Chapter 5 we evaluated the interval CRC incidence rate and FIT sensitivity after the first screening round and the impact of the adjustment of FIT cut-off. We observed a low cumulative incidence of interval CRCs because of the high sensitivity of FIT for CRC. We also concluded that there is an optimum in FIT cut-off at which it is not beneficial (i.e. lowering referrals for colonoscopy with restricted resources) to further increase the FIT cut-off. The main reason for this is that above 80 µg Hb/g faeces there is a large decrease in FIT sensitivity for CRC, while decrease in positivity rate (i.e. number of referrals) is mild. This mild decrease in positivity rate at higher cut-offs was also observed in the FIT pilot study in England.[24] The current FIT cut-off in the Dutch CRC screening programme is far below 80 µg Hb/g faeces. This again confirms that the increase to 47 µg Hb/g faeces was a prudent decision to optimise the screening programme. The outcomes of the study in Chapter 5 however may be informative to design a more tailored screening strategy. Using the information of previous FIT result, individuals at highest risk for interval CRC can be identified. Recent analysis of our research group has shown that individuals with an Hb concentration just below the cut-off of 47 µg Hb/g faeces were 16 times more likely to have an interval CRC. Consequently, tailored screening intervals could be designed to increase the benefits of screening while reducing the harms.

Socioeconomic differences

In **Chapter 6** we evaluated differences in FIT screening by social economic status (SES). We used area SES and compared the performance indicators participation rate, positivity rate, PPV and detection rate for CRC and AA between different SES quantiles. We concluded that CRC and AA yield per invitee does not differ by SES in the Dutch CRC screening programme. FIT screening even has the potential to reduce health inequalities in CRC mortality, because of a higher yield in participants with the lowest SES. However, this is currently offset by the lower participation in this group. A recent review confirms this variation in participation across SES in organised programmes worldwide.[25] Targeting individuals with the lowest SES could be beneficial, as highest health gains can be achieved in this group. However, similar to participation in general, individuals' motive for nonparticipation is unknown. Therefore, it is not clear what the best method is to inform and motivate individuals with the lowest SES to participate in screening. In England extensive research has been conducted to assess the impact of different evidence-based interventions on participation among individuals with the lowest SES. All different types of written materials enclosed with the invitation had no impact on the participation rate; GIST (Goals, Ideas, Step-projects, and Tasks)-based leaflet, narrative leaflet or GP endorsed

invitation.[26] The only method that showed a small increase in participation rate for individuals with the lowest SES was an enhanced reminder letter.[26] Not only an association with SES was observed, but the overall participation increased with an enhanced reminder letter.

Consistency of FIT performance

In the study in **Chapter 7** we estimated consistency of FIT performance over time on positivity and detection rates of CRC and AA within a national FIT-based CRC screening programme. Variation was observed for FIT positivity rate and detection rates of CRC and AA between FIT specimen collection devices (batches) and reagent lots, but no difference in PPV was identified. Based on these outcomes we concluded that clinically the programme is performing well, but there is room for improvement of the current quality assessment of the FIT within the Netherlands. Currently, no acceptable ranges of variation in positivity rate or Hb concentration exist. As a consequence of the observed variation, a discussion was started among parties involved in the Dutch CRC screening programme. Surprisingly, the discussion was not on the result itself, but more on the current set-up of the Dutch quality assurance system for FIT screening. It was realised that there were no acceptable range for observed variation and how to deal with observed variation. A discussion was initiated on important quality aspect of the performance of the FIT test including acceptable range of variation and value assignment. Daily controls in the participating laboratories are currently supervised by the Dutch Foundation for Quality Assessment in Medical Laboratories (SKML). They are carrying out three groups of controls, as described in detail in Chapter 7. However, international standardisation of quality assessment is lacking. Consequently, a national working group 'Quality assurance FIT' was set-up. The aim of this working group is to improve quality assessment of FIT screening in the Netherlands, combining the expertise of different stakeholders.

9.2 LIMITATIONS / METHODOLOGICAL CONSIDERATIONS

The most important event during the implementation phase of the Dutch CRC screening programme was the adjustment of FIT cut-off in the first year. As a consequence, we were able to compare important performance indicators by two cut-offs. A limitation of using data from the implementation phase of the Dutch CRC screening programme is that the screening programme was implemented by birth cohort. The conclusions in this thesis could therefore only be based on selected age groups and not on all age groups (55-75 years old) within the target population. Moreover, for the comparison of different rates of performance indicators by two FIT cut-offs, age-adjusted rates had to be calculated using direct standardisation. Because of large differences in the sample size per age group, wide confidence intervals (CI) were observed. Based on these wide CI, we concluded, for example for FIT sensitivity, that there was no significant difference between the different FIT cut-offs. However, the wide CI

indicates the uncertainty of the age-adjusted rates. Therefore we subsequently carried a logistic regression analysis, which showed a significant difference between the two cut-offs. Another limitation is that we started with older age groups; therefore the conclusions of this thesis are predominantly based on older age groups. In 2019, all age groups will have been invited at least once. More certainty about the performance indicators of the Dutch CRC screening programme should be obtained in the coming years, when all age groups have been invited, to ensure the programme is still performing in line with the expectations.

Another limitation is the completeness of the national information system (ScreenIT). ScreenIT is an excellent system which automates a large part of the screening process, from selection of eligible individuals of screening to obtaining result of endoscopy. However, in the first place it has been developed to structure the logistics of the screening process. As a result, obtaining data for scientific purpose has been a challenge. For example, for each individual polyp detected, pathology results cannot be linked directly to endoscopy results yet. This means that we know for a specific individual from the endoscopy report that there were four polyps detected and removed and one of them was large (> 10 mm). From the pathology report, we know that three of the polyps were adenomas and one was hyperplastic. We cannot distinguish whether the large polyp from the endoscopy was the adenoma or hyperplastic. In such situations we assumed the large polyp to be adenomatous, which is reasonable in most cases, but may have led to a small overestimation of AA. Another problem is that, especially as a result of long waiting periods in the first year, individuals may have had a colonoscopy scheduled outside of the screening programme. The results of these colonoscopies were not entered in ScreenIT. We estimated with data of national pathology database PALGA that 6-8% of the individuals with a positive FIT in 2014 had a colonoscopy outside the programme. It is expected that this percentage will be lower now the programme is fully implemented and waiting periods have been reduced and more colonoscopy centres joined the screening programme.

The last limitation is the difference in data collection and definition of performance indicators between the four CRC screening programmes described in Chapter 8. In France there is no national centralised information system collecting data on quality indicators and diagnostic yield. In Flanders a different definition of AA was used, because no data on size of the polyps was available. These limitations made it difficult to compare the CRC and AA yield of the four CRC screening programmes. This emphasises the importance of standardisation of important performance indicators for international comparison purposes.

9.3 FUTURE PERSPECTIVES

At the end of the implementation phase we may conclude that the Dutch CRC screening programme has reached a steady state and is performing in line with expectations. The main

goal should be to maintain this stable CRC programme with optimal programme performance. Nevertheless, there is always potential to further optimise a screening programme. The current programme could be expanded by inviting more age groups, a new test could be chosen or risk-factors could be used to invite those individuals' with highest CRC risk.

Expansion of the current screening programme

The Dutch CRC screening programme could be expanded in three ways: by lowering the starting age, lowering the FIT cut-off or adjusting the screening interval. Recently, the American Cancer Society changed their recommendations for CRC screening to start at a lower age (45 years old). The reason for this change is the observed increase in CRC incidence in young adults.[26] A similar increase has been observed in Europe.[28] Lowering the starting age will also be more in line with the European recommendation, that advices CRC screening in men and women aged 50-74.[29] However, it is unclear yet what the impact is of the increased incidence in young adults on the effectiveness of the Dutch CRC screening programme.

Rationale for lowering the FIT cut-off comes from a previous decision analysis. This analysis showed that increasing the FIT cut-off was the most effective option with limited colonoscopy resources compared to postpone screening in certain age groups.[10] The same could apply the other way around; lowering the FIT cut-off is the most effective option if the current programme can be expanded. The decision about whether lowering the FIT cut-off, widening the age range or shortening of the interval should be considered for expansion of the programme can be informed by model decision analyses. The outcomes of the studies included in this thesis can be used to inform several important decision model parameters.

New test modality

FIT has a high sensitivity for CRC, as also described in this thesis, but FIT has a lower sensitivity for AA (31%).[30] New test modalities are developed aiming for a higher sensitivity for CRC and AA. Stool DNA testing and video endoscopy seem to be promising new test modalities. They are not considered cost-effective yet, because their high costs do not outweigh the small additional benefit.[31] A new development is the use of protein biomarkers for the detection of CRC or AA, instead of or supplement to Hb protein. An early clinical phase of biomarker development study showed that new stool-based protein biomarkers have a higher discriminatory power than Hb protein alone.[30] A combination of 4 proteins resulted in a sensitivity for CRC of 80% similar to FIT and for AA of 45%, substantially higher than FIT. Potentially, these proteins could easily be implemented in a national FIT-based screening programme. Other promising new test modalities are FIT combined with blood markers for CRC or AA detection. A recent review concluded that most studies have not been able to prove improved FIT characteristics.[32] However, DNA hyper methylation markers seem most suitable, specifically methylated Septin 9 DNA plasma assay (^{m}SEPT9). This could be potentially a good alternative for CRC screening. However, it can only be cost-effective compared to

FIT-based screening if participation with FIT will drop below 70%.[33] The current participation rate in the Netherlands is still above 70%, therefore with the current test characteristics of *m*SEPT9 it is not a better option than FIT. When considering a new screening test it is important to realise that the new test should be relatively cheap to be a good candidate to introduce on population level. Decision analysis showed that costs of a new biomarker test should not exceed 7-fold the costs of FIT.[34] Another important aspect considering a new and more sensitive test for CRC is the associated specificity. If the new test has a lower specificity, this also should be taken into account when this test will be introduced on a population level. Lower specificity will lead to an increase in number of individuals with a false-positive result, resulting in more individuals undergoing an unnecessary colonoscopy.

Personalised screening strategy

Contemplating on the outcomes of the performance indicators described in this thesis, it seems that differential FIT cut-off by gender or previous screening result might lead to a more (cost-) effective screening programme. Decision analyses can be performed comparing such strategies to identify the optimal screening strategy for the current situation with uniform screening.[10] Other possible new strategies might be applying different FIT cut-offs for the first and subsequent screening rounds or allocating different intervals based on previous Hb concentration.

The first option for a more personalised screening strategy is to screen men and women differently. Men have higher CRC and AA detection rates and higher FIT sensitivity for CRC than women.[35] A tailored screening strategy could focus on similar sensitivity for both men and women by lowering the cut-off for women. The other way around is also an option, aiming for similar PPV for both men and women by lowering the cut-off for men.[36] However, long-term effectiveness of screening in terms of life years gained, estimated with decision modelling, it shows similar effectiveness for men and women. This is mainly the result of the fact that women will have a longer life expectancy than men. This is in line with previous results of our research group, showing that a different screening strategy for men and women is not cost-effective.[37] Note, this decision analysis assumed unlimited colonoscopy capacity. If programmes have limited colonoscopy resources, applying different screening strategies for men and women may be cost-effective, but this need to be further studied.

Faecal Hb concentration in previous screening round is another risk factor for the development of CRC and AA that also can be used for personalised screening. Individuals with a faecal Hb concentration below the cut-off (between 15 and 47 µg Hb/g faeces) are at higher risk for the detection of CRC or AA during consecutive screening rounds than individuals without any faecal Hb detected.[38-41] The biological hypothesis behind this finding is that adenomas will progressively bleed when developing to carcinoma, and therefore even low concentrations of Hb may be an indication of the presence of adenoma.[42] Therefore, individuals with high Hb concentration in the previous screening round may benefit from

shorter screening intervals. Contrary, individuals without detectable Hb in the previous screening round may benefit from extended screening intervals. Personalised screening based on faecal Hb concentration has two advantages over other known risk factors like smoking, obesity, food intake or family history. The estimated hazard ratios for individuals with small amount of faecal Hb concentration compared to those with no detectable faecal Hb concentration are considerably higher than those reported for e.g. lifestyle or family history.[41] Another advantage of using faecal Hb concentration of the previous screening round is its availability. This information is already being registered in the national information system (ScreenIT). Therefore, additional questionnaires on obtaining information on lifestyle and family history, which could jeopardise screening participation, are not needed.

Ideally, all potential risk factors like gender, age and Hb concentration of the previous screening round will be combined in one prediction model. All these separate risk factors contribute to individual's risk of having a CRC or AA. All risk factors should be combined to determine a person's individual risk. A Flemish study showed that men aged 74 with Hb concentration of >200 µg Hb/g faeces were 58 times more likely to be diagnosed with a CRC than women aged 56 with Hb concentration of 15 µg Hb/g faeces.[43] It is unknown to what extent the complexity of personalised strategies will impact the adherence to screening. For genetic testing it is known that knowing your gene-based risk profile will increase the willingness to participate in screening with 43%.[44] It is unknown whether this increase in adherence would also hold for an approach using relatively simple risk factors like gender, age and faecal Hb concentration. It will be totally different from genetic testing and might not lead to the same increase. The overall participation in the Netherlands is already high, so we will be more concerned if personalised screening will negatively impact the high participation.

9.4 FINAL CONCLUSIONS AND RECOMMENDATIONS

From the results of the studies that are presented in this thesis, the following conclusions can be drawn:

- Piloting, planning and implementing of the Dutch CRC screening programme may serve as a best practice for many screening initiatives currently being organised worldwide.
- The FIT participation rate in the first screening round in the Netherlands (71.8%) was one of the highest across the world and remained stable in the second screening round.
- Participation in follow-up colonoscopy (77.8%) was short of the minimally acceptable level of 85% and lower than surrounding countries in Western-Europe.
- Adjustments of the cut-off from 15 to 47 µg Hb/ g faeces halfway through the first year of the programme was necessary to ensure that the programme met the intended balance of harms and benefits of CRC screening.

- Using a higher FIT cut-off (47 µg Hb/ g) had limited impact on the cumulative CRC and AA detection because a substantial part of the missed lesions was detected in the second screening round.
- There is a strong correlation between faecal Hb concentration and detection of CRC and AA in subsequent screening. Individuals with a faecal Hb concentration just below the current cut-off (15-47 µg Hb/g faeces) were 23 times more likely to have CRC or AA detected at subsequent screening than those without detectable faecal Hb.
- Screen-detected CRCs have more often a favourable stage distribution (stage I and II) (67%) than symptom-detected CRCs (40%). Screen-detected CRCs were more often located in the left colon and rectum (73%) than symptom-detected CRCs (65%).
- FIT showed a high sensitivity for CRC (85.5%) with an associated low cumulative incidence of interval CRCs.
- FIT screening could potentially have a higher yield in participants with the lowest SES, but this higher yield is currently offset by the lower participation in this group.
- The overall population-impact of the variations in FIT positivity and detection rates between specimen collection devices and reagent lots is expected to be modest, but there is room for improvement of quality assessment. Currently, acceptable ranges of variation are lacking.

Based on these conclusions, we formulated the following recommendations:

- Coming years CRC-related mortality rates need to be closely monitored to ensure that CRC-related mortality is indeed decreasing as a consequence of the introduction of CRC screening in the Netherlands.
- Given the low participation to follow-up colonoscopy (77.8%), future research should be undertaken to identify reasons for nonparticipation and options to increase colonoscopy participation rate to the recommend level of 85%.
- The observed variation in FIT performance between batches and lot reagents described in this thesis can be used as input for the international initiative for standardising FIT quality assessment and for improving a regular monitoring system to reduce the impact of test variation on detection of CRC and AA.
- Outcomes during the implementation phase are mainly based on older age groups. Close monitoring of participation rate, positivity rate and detection of CRC and AA is needed to obtain estimates for all age groups of the target population.
- Personalised screening based on previous faecal Hb concentration is an important next step to explore for further optimisation of the Dutch CRC screening programme. Future studies are needed that evaluate the effectiveness of applying different screening intervals based on previous faecal Hb concentration.

REFERENCES

1. Klabunde C, Blom J, Bulliard JL, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. J Med Screen. 2015;22(3):119-26.
2. Landelijke Monitoring Bevolkingsonderzoek Borstkanker: https://www.rivm.nl/nieuws/monitor-2014. [Cited 2019 April 4].
3. Landelijke Monitoring Bevolkingsonderzoek Baarmoederhalskanker: https://www.rivm.nl/documenten/landelijke-evaluatie-van-bevolkingsonderzoek-baarmoederhalskanker-leba-tm-2014. [Cited 2019 April 4].
4. 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.
5. 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.
6. 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.
7. van der Vlugt M, Grobbee EJ, Bossuyt PM, et al. Adherence to colorectal cancer screening: four rounds of faecal immunochemical test-based screening. Br J Cancer. 2017;116(1):44-9.
8. Denters MJ, Deutekom M, Bossuyt PM, et al. Involvement of previous non-participants cannot fully compensate for lower participation in a second round of FIT-screening. Cancer Epidemiol. 2013;37(3):330-5.
9. van Dam L, Korfage IJ, Kuipers EJ, et al. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? Eur J Cancer. 2013;49(10):2321-30.
10. 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.
11. de Klerk CM, Wieten E, Lansdorp-Vogelaar I, et al. Performance of two faecal immunochemical tests for the detection of advanced neoplasia at different positivity thresholds: a cross-sectional study of the Dutch national colorectal cancer screening programme. Lancet Gastroenterol Hepatol. 2019;4(2):111-8.
12. Doubeni CA, Jensen CD, Fedewa SA, et al. Fecal Immunochemical Test (FIT) for Colon Cancer Screening: Variable Performance with Ambient Temperature. J Am Board Fam Med. 2016;29(6):672-81.
13. Symonds EL, Osborne JM, Cole SR, et al. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. J Med Screen. 2015;22(4):187-93.
14. Dancourt V, Hamza S, Manfredi S, et al. Influence of sample return time and ambient temperature on the performance of an immunochemical faecal occult blood test with a new buffer for colorectal cancer screening. Eur J Cancer Prev. 2016;25(2):109-14.
15. Grazzini G, Ventura L, Rubeca T, et al. Impact of a new sampling buffer on faecal haemoglobin stability in a colorectal cancer screening programme by the faecal immunochemical test. Eur J Cancer Prev. 2017;26(4):285-91.
16. Moss S, Ancelle-Park R, Brenner H, International Agency for Research on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Evaluation and interpretation

of screening outcomes. Endoscopy. 2012;44 Suppl 3:SE49-64.

17. Doubeni CA, Fedewa SA, Levin TR, et al. Modifiable Failures in the Colorectal Cancer Screening Process and Their Association With Risk of Death. Gastroenterology. 2019;156(1):63-74 e6.
18. Elferink MAG, Toes-Zoutendijk E, Vink GR, et al. National population screening for colorectal carcinoma in the Netherlands: results of the first years since the implementation in 2014. Landelijk bevolkingsonderzoek naar colorectaal carcinoom. Ned Tijdschr Geneeskd. 2018;162:D2283.
19. Brenner H, Niedermaier T, Chen H. Strong subsite-specific variation in detecting advanced adenomas by fecal immunochemical testing for hemoglobin. Int J Cancer. 2017;140(9):2015-22.
20. de Wijkerslooth TR, Stoop EM, Bossuyt PM, et al. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia. Am J Gastroenterol. 2012;107(10):1570-8.
21. 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.
22. Chang LC, Shun CT, Hsu WF, et al. Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. Clin Gastroenterol Hepatol. 2017;15(6):872-9 e1.
23. JE IJspeert, de Wit K, van der Vlugt M, et al. Prevalence, distribution and risk of sessile serrated adenomas/polyps at a center with a high adenoma detection rate and experienced pathologists. Endoscopy. 2016;48(8):740-6.
24. 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. 2017;66(9):1631-44.
25. de Klerk CM, Gupta S, Dekker E, et al. Expert Working Group 'Coalition to reduce inequities in colorectal cancer screening' of the World Endoscopy O. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. GUT. 2018;67(4):679-87.
26. Wardle J, von Wagner C, Kralj-Hans I, et al. Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the English NHS Bowel Cancer Screening Programme (ASCEND): four cluster-randomised controlled trials. Lancet. 2016;387(10020):751-9.
27. Peterse EFP, Meester RGS, Siegel RL, et al. The impact of the rising colorectal cancer incidence in young adults on the optimal age to start screening: Microsimulation analysis I to inform the American Cancer Society colorectal cancer screening guideline. Cancer. 2018;124(14):2964-73.
28. Vuik FER, Nieuwenburg SAV, Bardou M, et al. Increasing incidence of colorectal cancer in young aduls in Europe over the last 25 years. GUT. 2019.
29. Council of the European Union. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). Off J Eur Union. 2003:34–38.
30. Bosch LJW, de Wit M, Pham TV, et al. Novel Stool-Based Protein Biomarkers for Improved Colorectal Cancer Screening: A Case-Control Study. Ann Intern Med. 2017;167(12):855-66.
31. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening - an overview. Best Pract Res Clin Gastroenterol. 2010;24(4):439-49.
32. Niedermaier T, Weigl K, Hoffmeister M, et al. Fecal immunochemical tests in combination with blood tests for colorectal cancer and advanced adenoma detection-systematic review. United European Gastroenterol J. 2018;6(1):13-21.

33. Ladabaum U, Allen J, Wandell M, et al. Colorectal cancer screening with blood-based biomarkers: cost-effectiveness of methylated septin 9 DNA versus current strategies. Cancer Epidemiol Biomarkers Prev. 2013;22(9):1567-76.
34. Lansdorp-Vogelaar I, Goede SL, Bosch LJW, et al. Cost-effectiveness of High-performance Biomarker Tests vs Fecal Immunochemical Test for Noninvasive Colorectal Cancer Screening. Clin Gastroenterol Hepatol. 2018;16(4):504-12 e11.
35. Toes-Zoutendijk E, van Leerdam ME, Dekker E, et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels. Gastroenterology. 2017;152(4):767-75 e2.
36. Grobbee EJ, Wieten E, Hansen BE, et al. Fecal immunochemical test-based colorectal cancer screening: The gender dilemma. United European Gastroenterol J. 2017;5(3):448-54.
37. Meulen MPV, Kapidzic A, Leerdam MEV, et al. Do Men and Women Need to Be Screened Differently with Fecal Immunochemical Testing? A Cost-Effectiveness Analysis. Cancer Epidemiol Biomarkers Prev. 2017;26(8):1328-36.
38. 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.
39. 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.
40. 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.
41. Grobbee EJ, Schreuders EH, Hansen BE, et al. Association Between Concentrations of Hemoglobin Determined by Fecal Immunochemical Tests and Long-term Development of Advanced Colorectal Neoplasia. Gastroenterology. 2017;153(5):1251-9 e2.
42. 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;147(3):628-36 e1.
43. van de Veerdonk W, Van Hal G, Peeters M, et al. Risk stratification for colorectal neoplasia detection in the Flemish colorectal cancer screening programme. Cancer Epidemiol. 2018;56:90-6.
44. Ramsey S, Blough D, McDermott C, et al. Will knowledge of gene-based colorectal cancer disease risk influence quality of life and screening behavior? Findings from a population-based study. Public Health Genomics. 2010;13(1):1-12.

APPENDICES

SUMMARY

Worldwide, the incidence of colorectal cancer (CRC) is increasing due to ageing of the population, change in dietary habits and rise in risk factors like smoking, obesity, and lack of physical activity. It is expected that without interference the number of CRC cases in the Netherlands will increase from 13,000 to 17,000 persons per year by 2020.

Worldwide, many countries have implemented a CRC screening programme, with the aim to reduce the high incidence and mortality rates of CRC. Choosing the best screening strategy for the population is a complex process. When deciding on which test to use, several aspects should be taken into account: test sensitivity and specificity, participation, harms, resource capacity and costs. The Netherlands may well serve as an excellent example for how all these aspects of screening were weighed into the decision for the most optimal CRC screening method for the Dutch population. In the Netherlands, an extensive preparatory process has taken place before the implementation of the national population-based CRC screening programme.

In 2006, pilot studies were initiated to study the potential of a national CRC screening programme in the Netherlands. The aim of these Dutch pilot studies was to evaluate the most important aspects (i.e. participation, diagnostic yield and cost-effectiveness) of the most relevant screening methods: guaiac Faecal Occult Blood Test gFOBT), Faecal Immunochemical Test (FIT; with FIT cut-offs ranging from 10-40 µg Hb/g faeces), sigmoidoscopy, colonoscopy and computed tomographic colonography (CTC). FIT screening showed the highest participation rate up to 60-62% in the first round, compared to 47-50% for gFOBT, 32% for sigmoidoscopy, 34% for CTC and 22% for colonoscopy. FIT also showed the highest diagnostic yield, with the highest detection of CRC per 1,000 invitees over two screening rounds. Based on these outcomes, subsequent modelling and the advice of the Dutch Health council it was decided to implement FIT screening, with a cut-off of 15 µg Hb/g faeces for a positive result. This cut-off was predominantly preferred because of a more favourable balance between true-positives and false-positives (higher positive predictive value) than a cut-off of 10 µg Hb/g faeces. It would also result in a lower colonoscopy demand.

The Dutch CRC screening programme was gradually implemented by age group from 2014 onwards, with biennial for men and women aged 55 to 75 years. While the design of a CRC screening programme may be evidence-based and well-planned, this does not guarantee that the performance on a national level will be in line with the expectations. Therefore, monitoring a screening programme is essential, especially during the implementation phase.

To ensure that these expectations are met on a national level, monitoring and evaluation of the national screening programme in real setting and/or national level is important. The general aim of this thesis is therefore to evaluate performance of the Dutch colorectal cancer screening programme during the implementation phase. We evaluated the participation to FIT, FIT positivity, participation to follow-up colonoscopy, CRC and advanced adenoma (AA)

detection in the first screening round (Chapter 2) and second screening round (Chapter 4), the stage- and location distribution of screen-detected cancers (Chapter 3), and the incidence interval cancers (Chapter 5). For all these performance indicators, the impact of the adjusted cut-off in the first half year was evaluated. The most important performance indicators were also compared between different surrounding countries that implemented FIT screening (Chapter 8).

Participation FIT

High participation rate is a relevant finding, as high participation is an important indicator for a successful screening program. We observed high participation rate (71.3%) in the first year of the national CRC screening programme. This rate is one of the highest in the world and remained high in the second screening round (75.9%), driven by a consistent participation of 93% among the individuals that previously participated. Moreover, 21% of non-participants in the first round sent in a FIT in the second screening round.

FIT positivity

Real-time monitoring allowed for instant adjustment of the programme when the outcomes substantially differed from what was expected. Higher positivity rate (10.6% versus expected 6.4%) together with the lower than anticipated PPV for CRC and AA (42.1% versus expected 51.6%) in the first half year of implementation, resulted in the decision to increase the cut-off for a positive FIT from 15 to 47 μg Hb/g faeces. This adjustment resulted in a programme that more closely met the expectations, with a positivity rate of 6.7% and PPV for CRC and AA of 49.1%. The better balance of harms and benefits of the screening programme was also more in accordance with the recommendation of the Dutch Health Council before the start of the programme. As expected, the positivity rate dropped in the second screening round to 3.3% for those previously tested with the low cut-off and to 4.3% for those previously tested with the higher cut-off.

Participation to follow-up colonoscopy

The high FIT participation rate in the current Dutch programme did not apply to participation to follow-up colonoscopy. The observed participation to the colonoscopy follow-up after a positive FIT of 77.8% in the Netherlands is below the minimally acceptable recommended level of 85% and lower than surrounding countries (92.8% in the Basque country, 88.4% in France and 82.8% in Flanders). The rate in the Netherlands is most likely an underestimation, as some individuals (6-8%) may have had a colonoscopy in centres outside the screening programme. These colonoscopy results were not integrated in the national screening information system ScreenIT. Effort should be undertaken to increase this relatively low participation rate, because individuals without appropriate follow-up after a positive FIT are seven times more likely to die from CRC than individuals with appropriate follow-up.

Colorectal cancer and advanced adenoma detection

The PPV as observed in the first half year of 2014 of 42.1% was below the desired PPV of 51.6%. As a result of the increased FIT cut-off, the PPV for CRC and AA increased to a more desirable level of 49.1%. However, the detection rate decreased for CRC (5.8‰ to 4.4‰) and AA (30.8‰ to 20.6‰), indicating that CRCs will be missed in the first screening round. Based on the results of the second screening round we concluded that using a higher FIT cut-off has limited impact on the CRC and AA detection overall, because as a substantial part of the missed lesions will indeed be detected in subsequent screening rounds. However, after two screening rounds the cumulative yield of CRC and AA is lower for those tested with the higher cut-off in the first year compared to those tested with the lower cut-off in the first year. We also showed that individuals with a negative FIT in the first round, but with a faecal Hb concentration between 15-47 µg Hb/g faeces (just below FIT cut-off) had 23 times the odds of having CRC or AA detected in the consecutive screening round than individuals with no detectable faecal Hb.

Stage distribution and location

Screen-detected CRCs showed a more favourable stage distribution (stage I and II, 66.7%) than symptom-detected CRCs (stage I and II, 39.8%). These findings are very encouraging and it is a promising sign that screening may decrease CRC-related morbidity and eventually mortality rates. We also showed that screen-detected CRCs were more often located in the left hemicolon than symptom-detected CRCs (45.9% versus 31.4%). Combining this with the outcomes of interval CRCs we conclude that FIT has a lower sensitivity for right-sided cancers, as many more interval cancers were located right-sided. The first explanation for this finding might be that FIT is less sensitive for right-sided cancers due to degradation of Hb during colon transit. Another explanation might be that FIT is less sensitive for sessile serrated polyps and this type of polyps is more often detected in the right colon.

Interval cancers

We observed a low two-year cumulative incidence of interval CRCs after negative FIT (11.2 per 10,000 individuals) because of the high sensitivity of FIT for CRC (85.5%). The age-adjusted cumulative incidence of interval CRCs of 9.5 per 10,000 individuals at the low cut-off was significantly lower than the 13.8 per 10,000 individuals at the higher cut-off. Age-adjusted FIT sensitivities for CRC were statistically different with 90.5% at the low cut-off and 82.9% at the higher cut-off. FIT sensitivity for CRC was higher for men (87.4%) than for women (82.6%).

Socioeconomic differences

Screening performance indicators were also compared by socioeconomic status (SES), with Quintile 1 being the least deprived and Quintile 5 the most deprived. The analyses showed

that performance indicators vary by SES. Participation to FIT screening was significantly lower for individuals in the lowest SES quintile (67.0%) compared to the other four quintiles (73.0% to 75.1%). A similar pattern was observed for colonoscopy participation after a positive FIT: 75.8% for individuals in the lowest SES quintile and 80.0-82.4% for individuals in the other quintiles. On the other hand, the detection rate per FIT participant for advanced neoplasia gradually increased from 3.3% in Quintile 1 to 4.0% in Quintile 5. The lower participation but higher yield per participant resulted in a similar yield per invitee for the lowest (2.04%) and the highest SES quintile (2.00%). These results show that screening has the potential to reduce health inequalities in CRC mortality, because of a higher detection in more deprived participants. However, this is currently offset by the lower participation in this group.

Consistency of FIT performance

We observed small variation in FIT positivity rate and detection rates of CRC and AA between FIT specimen collection devices (batches) and reagent lots. In contrast, PPV was not found to vary between collection devices, reagents or laboratories. The size of the population that was impacted by these deviating rates was small (0.1% of the total tested population). Based on these outcomes we concluded that clinically the Dutch CRC screening programme is performing well, but there is room for improvement. Currently, no recommendation on the variation in positivity rate or Hb concentration exists. Effort is undertaken in the Netherlands to improve the quality assessment of the FIT within the screening programme.

Conclusions and recommendations

From the results of the studies that are presented in this thesis, the following conclusions can be drawn:

- Piloting, planning and implementing of the Dutch CRC screening programme may serve as a best practice for many screening initiatives currently being organised worldwide.
- The FIT participation rate in the first screening round in the Netherlands (71.8%) was one of the highest in the world and remained stable in the second screening round.
- Participation in follow-up colonoscopy (77.8%) was below the minimally acceptable level of 85% and lower than surrounding countries in Western-Europe.
- Adjustments of the cut-off from 15 to 47 µg Hb/ g faeces halfway through the first year of the programme was necessary to ensure that the programme met the intended balance of harms and benefits of CRC screening.
- Using a higher FIT cut-off (47 µg Hb/ g) had a limited impact on the cumulative CRC and AA detection because a substantial part of the missed lesions was detected in the second screening round.
- There is a strong correlation between faecal Hb concentration and detection of CRC and AA in subsequent screening. Individuals with a faecal Hb concentration just below the

current cut-off (15-47 µg Hb/g faeces) were 23 times more likely to have CRC or AA detected at subsequent screening than those without detectable faecal Hb.
- Screen-detected CRCs more often have a favourable stage distribution (stage I and II) (67%) than symptom-detected CRCs (40%). Screen-detected CRCs were more often located in the left colon and rectum (73%) than symptom-detected CRCs (65%).
- FIT showed a high sensitivity for CRC (85.5%) with an associated low cumulative incidence of interval CRCs.
- FIT screening could potentially have a higher yield in participants with the lowest SES, but this higher yield is currently offset by the lower participation in this group.
- The overall population-impact of the variations in FIT positivity and detection rates between specimen collection devices and reagent lots is expected to be modest, but there is room for improvement of quality assessment. Currently, acceptable ranges of variation are lacking.

Based on these conclusions, we formulated the following recommendations:
- In the coming years, CRC-related mortality rates need to be closely monitored to ensure that CRC-related mortality is indeed decreasing as a consequence of the introduction of CRC screening in the Netherlands.
- Given the low participation to follow-up colonoscopy (77.8%), future research should be undertaken to identify reasons for nonparticipation and options to increase colonoscopy participation rate to the recommend level of 85%.
- The observed variation in FIT performance between batches and lot reagents described in this thesis can be used as input for the international initiative for standardising FIT quality assessment and for improving a regular monitoring system to reduce the impact of test variation on detection of CRC and AA.
- Outcomes during the implementation phase are mainly based on older age groups. Close monitoring of participation rate, positivity rate and detection of CRC and AA is needed to obtain estimates for all age groups of the target population.
- Personalised screening based on previous faecal Hb concentration is an important next step to explore for further optimisation of the Dutch CRC screening programme. Future studies are needed that evaluate the effectiveness of applying different screening intervals based on previous faecal Hb concentration.

SAMENVATTING

Wereldwijd neemt de incidentie van darmkanker toe als gevolg van de vergrijzing van de bevolking, verandering van eetgewoontes en ongezonde leefstijl zoals roken, obesitas en onvoldoende beweging. De verwachting is dat zonder interventie het aantal nieuwe diagnoses van darmkanker in Nederland zal stijgen van 13.000 nu naar 17.000 gevallen per jaar in 2020.

Verschillende landen hebben in de afgelopen jaren een bevolkingsonderzoek naar darmkanker geïmplementeerd, met als doel de incidentie en sterfte ten gevolge van darmkanker te verminderen. De keuze voor de beste strategie om te screenen is complex. Verschillende aspecten moeten overwogen worden; sensitiviteit, specificiteit, deelname, nadelen (bijvoorbeeld complicaties), capaciteit en kosten. Nederland is een goed voorbeeld van een land waarbij deze verschillende aspecten zijn afgewogen voordat er een beslissing is genomen over de meest optimale opzet voor het Nederlandse bevolkingsonderzoek darmkanker.

In 2006 zijn er proefbevolkingsonderzoeken gestart om de mogelijkheden voor invoering van een landelijk bevolkingsonderzoek in kaart te brengen. De belangrijkste uitkomsten zoals deelname, detectiecijfer darmkanker en advanced adenomen (AA) en kosten-effectiviteit zijn vergeleken voor verschillende screeningtesten: de ontlastingstesten guaiac faeces occult bloed test (gFOBT) en faeces immunochemische test op occult bloed (FIT), sigmoïdoscopie, coloscopie, en CT-colonografie. Met een deelnamepercentage van 60-62% was deelname het hoogst voor genodigden met FIT, gevolgd door 47-50% voor gFOBT, 32% voor sigmoïdoscopie, 34% voor CT-colonografie en 22% voor coloscopie. Per 1.000 genodigden werden met de FIT ook deze meeste darmkankers gevonden. Op basis van deze uitkomsten, samen met de uitkomsten van kosten-effectiviteitanalyse en advies van de gezondheidsraad, is besloten om FIT met een afkapwaarde van 15 µg Hb/g ontlasting in te voeren voor het landelijke bevolkingsonderzoek. Deze beslissing was met name gebaseerd op het streven naar een goede balans tussen de voordelen (terecht-positieven, darmkanker detectie) en nadelen (fout-positieven, onnodige coloscopie) van het bevolkingsonderzoek.

Het bevolkingsonderzoek darmkanker werd van 2014-2018 per geboortecohort gefaseerd ingevoerd. De uiteindelijke doelgroep van het bevolkingsonderzoek darmkanker bestaat uit mannen en vrouwen van 55 tot en met 75 jaar, die elke twee jaar uitgenodigd worden voor screening d.m.v. de (FIT). Monitoren van een bevolkingsonderzoek is belangrijk, met name gedurende de invoeringsfase van een nieuw programma. De opzet van het programma kan goed onderzocht en gepland zijn, maar het is onbekend of de prestatie van het programma op landelijke niveau overeen zal komen met de verwachting.

Na jaren van voorbereiding, waren de verwachtingen van het landelijke bevolkingsonderzoek hoog. Het doel van dit proefschrift is daarom om de prestatie van het landelijke bevolkingsonderzoek te evalueren gedurende de implementatie fase. De

belangrijkste prestatie indicatoren, zoals deelnamecijfer, verwijscijfer, deelnamecijfer coloscopie en detectiecijfer darmkanker en AA, werden geëvalueerd voor eerste ronde genodigden (hoofdstuk 2) en tweede ronde genodigden (hoofdstuk 4). Daarnaast hebben we de stadiumverdeling van screen-gedetecteerde darmkankers (hoofdstuk 3) en het aantal interval kankers na negatieve FIT (hoofdstuk 5) geanalyseerd. Voor bovenstaande uitkomsten hebben we ook de impact van het verhogen van de afkapwaarde halverwege het eerste jaar berekend. Als laatste hebben we de uitkomsten van het Nederlandse programma vergeleken met programma's in omliggende landen: Baskenland (Spanje), Frankrijk en Vlaanderen (België) (hoofdstuk 8).

Deelnamecijfer

In het eerste jaar van het Nederlandse bevolkingsonderzoek darmkanker was de deelname hoog; 71.3%. Daarmee heeft Nederland de hoogste opkomst wereldwijd. De uitkomsten van de tweede uitnodigingsronde lieten zien dat de deelname met 75.9% hoog blijft in vervolgrondes; 93% van de deelnemers in de eerste ronde neemt opnieuw deel en 21% van de niet-deelnemers in de eerste ronde neemt nu wel deel.

FIT verwijscijfer

Het wekelijks monitoren van de uitkomsten van het bevolkingsonderzoek liet zien dat het programma niet presteerde in lijn met onze verwachting waardoor een aanpassing noodzakelijk was. In het eerste half jaar van het bevolkingsonderzoek was er namelijk een hoger dan verwacht verwijscijfer (10,6% versus de verwachtte 6,4%) met een lager dan verwacht positief voorspellende waarde voor darmkanker en AA (42,1% versus de verwachtte 51,6%). Dit heeft er toe geleid dat de afkapwaarde voor een positieve test verhoogd is van 15 naar 47 µg Hb/g ontlasting, met als resultaat een lager verwijscijfer (6,7%) en hogere positief voorspellende waarde (49,1%). Hiermee voldoet de prestatie van het programma meer aan de verwachting, met een betere balans tussen de voordelen en nadelen van het bevolkingsonderzoek darmkanker. Zoals verwacht daalde het verwijscijfer in de tweede uitnodigingsronde ten opzichte van de eerste ronde naar 3,3% voor personen die eerder getest waren met de lage afkapwaarde en 4,3% voor personen die eerder getest waren met de hoge afkapwaarde.

Deelname aan coloscopie

Deelnemers met een positieve FIT uitslag worden uitgenodigd voor een vervolgcoloscopie. Deelname aan deze coloscopie is met 77,8% lager dan de in Europa geadviseerde 85% en is ook lager in vergelijking met omliggende landen (deelname van 92,8% in Baskenland, 88,4% in Frankrijk en 82,8 in Vlaanderen). Het deelnamecijfer in Nederland is hoogstwaarschijnlijk een onderschatting omdat 6-8% van de deelnemers buiten het georganiseerde programma een coloscopie ondergaat. De bevindingen bij deze coloscopieën komen niet in de

landelijke screeningdatabase ScreenIT. Desalniettemin is het belangrijk dat er actie wordt ondernomen om dit lage deelnamecijfer te verhogen. Personen met een positieve FIT zonder vervolgcoloscopie hebben namelijk zeven maal meer kans om te overlijden aan darmkanker dan personen die wel zijn gekomen. Grotere betrokkenheid van huisartsen, die het belang van het vervolgonderzoek kunnen toelichten, zou eventueel tot een stijging in de deelname aan coloscopie kunnen leiden.

Detectiecijfer van darmkanker en advanced adenomen

Zoals reeds hierboven vermeld, was de positief voorspellende waarde in het eerste half jaar lager dan gewenst: 42,1% versus 51,6%. Een direct effect van het verhogen van de afkapwaarde was een hogere positief voorspellende waarde voor darmkanker en AA, namelijk 49,1%. De verhoging van de afkapwaarde had ook een negatief gevolg. Het detectiecijfer voor darmkanker daalde van 5,8‰ naar 4,4‰ en voor AA daalde het van 30.8‰ naar 20.6‰. Dit wijst er op dat er in de eerste ronde kankers zijn gemist. In hoofdstuk 4 concluderen we echter dat de verhoging van de afkapwaarde bij herhaalde deelname minimale impact heeft op het aantal gevonden darmkankers en AA. Een groot deel van de gemiste laesies wordt namelijk in de tweede uitnodigingsronde gevonden. Ondanks dat het meerendeel van de gemiste laesies in de volgende ronde wordt gevonden, was de cumulatieve opbrengst van darmkanker en AA over twee uitnodigingsrondes hoger voor personen die in de eerste ronde met een lage afkapwaarde waren getest. We toonden ook aan dat personen met een negatieve FIT in de eerste ronde, maar met een hemoglobine level net onder de huidige afkapwaarde (FIT uitslag tussen 15 en 47 µg Hb/g ontlasting), 23 keer meer kans hebben op het ontdekken een darmkanker of AA in de vervolgronde in vergelijking met personen zonder bloed in de ontlasting.

Stadiumverdeling en locatie

Screen-gedetecteerde darmkankers hebben een gunstigere stadiumverdeling (stadium I en II, 66,7%) in vergelijking met symptoom-gedetecteerde darmkankers (39,8%). Deze resultaten zijn veelbelovend en lijken erop te duiden dat het bevolkingsonderzoek darmkanker zal leiden tot een afname in darmkanker gerelateerde ziekte en sterfte. We constateerden ook dat screen-gedetecteerde kankers vaker linkszijdig worden gevonden ten opzichte van symptoom-gedetecteerde damkankers. Op basis van deze resultaten kunnen we concluderen dat FIT een lagere sensitiviteit heeft voor rechtszijdige kankers, wat bevestigd werd door de uitkomsten van de interval kankers. Een mogelijke verklaring hiervoor is dat de FIT minder gevoelig is voor rechtszijdige kankers als gevolg van degradatie van het hemoglobine tijdens passage door de darm. Een andere verklaring is dat FIT minder gevoelig is voor sessile serrated poliepen en dit type poliep zit voornamelijk rechtszijdig in de darm.

Interval kankers

We toonden in hoofdstuk 5 aan dat de FIT een hoge sensitiviteit (85,5%) heeft voor darmkanker, met als gevolg een lage cumulatieve incidentie van interval kankers (11,2 per 10.000 personen). Cumulatieve incidentie van interval kankers was verschillend voor de twee afkapwaardes: personen getest met 47 µg Hb/g ontlasting hadden een iets hogere cumulatieve incidentie (13,8 per 10.000 personen) dan personen getest met 15 µg Hb/g ontlasting (9,5 per 10.000 personen). De FIT met een afkapwaarde van 15 µg Hb/g ontlasting heeft een sensitiviteit voor darmkanker van 90,5%. De sensitiviteit was lager bij een FIT afkapwaarde van 47 µg Hb/g (82,9%). FIT heeft een hogere sensitiviteit in mannen (87,4%) dan in vrouwen (82,6%)

Sociaal economische verschillen

We hebben ook de prestatie indicatoren vergeleken voor personen met een verschillende sociaal economische status (SES), opgedeeld in vijf groepen. Deelname aan FIT was lager voor personen met de laagste SES (67,0%) in vergelijking met de andere vier SES groepen (73,0-75,1%). Een soortgelijk patroon was te zien voor deelname aan de coloscopie, met de laagste deelname voor personen met de laagste SES (75,8%) in vergelijking met de andere vier SES groepen (80,0-82,4%). Personen met de laagste SES hadden vaker een positieve FIT en ook het detectiecijfer van darmkanker en AA per deelnemer was hoger bij personen in een lagere SES groep (hoogste SES 3,3% en laagste SES 4,0%). Als gevolg van een lagere deelname aan de coloscopie maar een hoger detectiecijfer van darmkanker en AA in de lagere SES, waren er geen verschillen tussen de vijf SES groepen in het detectiecijfers voor darmkanker en AA per genodigden. FIT screening heeft de mogelijkheid om de gezondheidsongelijkheid ten gevolge van darmkanker sterfte te verminderen, door het hogere detectiecijfer van darmkanker en AA. Echter wordt dit nu gehinderd door een lagere deelname aan FIT screening door personen met de laagste SES status.

Consistentie FIT prestatie

We lieten kleine verschillen zien voor het verwijscijfer en het detectiecijfer van darmkanker en AA als personen werden getest met FIT buizen afkomstig uit verschillende leveringen of met verschillend reagens lots. Er werd hierin geen verschil waargenomen voor de positief voorspellende waarde voor darmkanker en AA. Deze verschillen hebben echter betrekking op een zeer kleine groep personen (0,1% van de totale groep). We concluderen daarom dat het programma klinisch goed presteert, maar dat er op analytisch gebied nog ruimte is voor verbetering. Op dit moment zijn er geen afspraken gemaakt over welke variatie in prestatie van FIT acceptabel is. Daarom is er een landelijke werkgroep opgericht, die zich bezig houdt met de kwaliteits monitoring van de FIT en de analyse daarvan.

Op basis van de resultaten gepresenteerd in dit proefschrift, kunnen de volgende conclusies worden getrokken:

- De implementatie van het Nederlandse bevolkingsonderzoek darmkanker dient als een goed voorbeeld voor veel andere programma's wereldwijd.
- De verhoging van de afkapwaarde halverwege het eerste jaar was noodzakelijk om weer een goede balans te vinden tussen de voordelen en nadelen van deelname aan het Nederlandse bevolkingsonderzoek darmkanker.
- Het deelnamecijfer aan FIT screening in Nederland is één van de hoogste wereldwijd en blijft hoog in vervolgrondes.
- Deelname aan de vervolgcoloscopie in Nederland is lager dan de geadviseerde 85% en ook lager dan omliggende landen.
- De hogere afkapwaarde heeft minimale impact op het detectiecijfer van darmkanker en AA, omdat een groot deel van de laesies in de vervolgronde wordt gedetecteerd.
- Er is een sterkte associatie tussen de hemoglobine concentratie in de ontlasting en het detectiecijfer van darmkanker of AA in een vervolgronde.
- Screen-gedetecteerde darmkankers hebben een gunstigere stadiumverdeling en worden vaker linkszijdig gevonden dan symptoom-gedetecteerde darmkankers.
- FIT heeft een hoge sensitiviteit voor darmkanker met een lage cumulatieve incidentie voor interval kankers.
- FIT detectiecijfer voor darmkanker en AA per genodigde zou hoger kunnen zijn voor personen met een lagere SES, maar wordt tenietgedaan door een lager deelnamecijfer.
- De impact van de variatie in verwijscijfer en detectiecijfer tussen verschillende leveringen van buizen en reagentia lots is klein, maar er is ruimte voor verbetering.

Op basis van deze conclusies kunnen de volgende aanbevelingen worden gedaan:

- De evaluatie van de prestatie-indicatoren zijn veelbelovend en het lijkt erop dat als gevolg van FIT screening de darmkanker gerelateerde ziekte en sterfte zal afnemen. Komende jaren zal de darmkanker gerelateerde stefte gemonitord moeten worden om vast te stellen of dit inderdaad het geval is.
- Vanwege de lage deelname aan de vervolgcoloscopie, zal er meer onderzoek gedaan moeten worden naar de reden van niet deelname en naar opties om deze deelname te verhogen.
- De geobseerveerde variatie tussen leveringen van buizen en reagentia lots kan gebruikt worden om internationaal standaarden vast te stellen voor kwaliteitscontrole van de FIT. Daarnaast kan het gebruikt worden om kort-cyclische monitoring op te zetten.
- Het bevolkingsonderzoek kan verder geoptimaliseerd worden door personen uit te nodigen op basis van het risico op darmkanker of AA door gebruik te maken van de hemoglobine concentratie in de voorgaande ronde.

DANKWOORD

Waarschijnlijk het best gelezen hoofdstuk van dit boekje, het dankwoord. Ik heb mijn best gedaan om het daarom zo lang mogelijk te maken.

Lieve Iris, zonder jou was er geen boekje. Toen ik jaren geleden kwam kennismaken omdat ik had gehoord dat er een vacature was, voelde ik me direct bij je op mijn gemak. Ik solliciteerde niet zozeer op het onderwerp, maar meer op jou als begeleider. Dit bleek een goede keuze, het was heel fijn om met jou samen te werken. Je weet ontzettend veel en bent een oprecht en vriendelijk persoon; negatieve feedback geven is zelfs een ontwikkelpunt van je ☺. Je geeft veel vrijheid en daarmee een gevoel van vertrouwen. Ik vond het lastig om bij MGZ weg te gaan, omdat ik het zo goed naar mijn zin heb gehad. Bedankt voor alles!

Monique, jou als tweede co-promotor was een perfecte aanvulling op Iris. Jouw gedrevenheid motiveerde mij om meer uit mezelf te halen. Ik waardeer je eerlijkheid en directe feedback, het was daardoor prettig samenwerken. Ondanks jouw drukke agenda blijf je attent op de kleine sociale dingen, bedankt daarvoor.

Harry, ik vond het leuk om bij jou te promoveren, met jouw jarenlange ervaring in screening heb ik veel van je mogen leren. De onderzoekslijn heeft een hoog niveau en het was een voorrecht daar deel van uit te maken. Over 1 ding zullen we het nooit eens worden: de figuren. Het moet je verbazen dat er zelfs een figuur op mijn cover staat. De (sectie)meetings, uitjes en borrels zijn een stuk gezelliger als jij erbij bent en jouw recept en restaurant tips zijn altijd goed!

Corine, jouw aanwezigheid als begeleider was zeer kort, maar toch wil ik jou hier noemen. Mijn start bij MGZ was flitsend. Ik kreeg een spoedcursus BI- en Excel-vaardigheden, zodat ik zelf de berekeningen voor de monitor kon doen, terwijl ik nog maar 4 maanden bij MGZ werkte. Dank daarvoor, want door jouw kennis en geduld heb ik het eerste jaar overleefd!

Lieve Annemieke, jij begon als begeleider van Arthur en mij in 2017. Tot groot verdriet liep het leven anders voor jou, waardoor je een tijd afwezig bent geweest. Toch heb je in het afgelopen jaar een belangrijke bijdrage geleverd aan onze artikelen en notities, jouw kritische blik en precisie zorgen voor betere stukken, dank je wel.

Professor Meijer, professor de Rijke en professor van Hal, hartelijk dank voor het beoordelen van mijn proefschrift en de bereidheid om plaats te nemen in de kleine commissie. *Professor Timmermans* and *Professor de Beaufort* dank voor het zitting nemen in de promotiecommissie.

LECO teamleden: Manon Spaander, Evelien Dekker, Ernst Kuipers, Iris Nagtegaal, Hanneke van Vuuren, Hans Bonfrer, Chris Ramakers, Katharina Biermann, Loes Velthuysen, Folkert van Kemenade, Maarten Thomeer, Harriët van Veldhuizen en Marjolein van Ballegooijen. Nu mijn promotie ten einde loopt besef ik me wat een voorrecht het is om als eerste op de data van het landelijke bevolkingsonderzoek darmkanker te promoveren. Velen van jullie hadden jaren voorbereiding achter de rug en ik was de geluksvogel om mijn naam op de papers te mogen zetten. Ondanks al jullie drukke agenda's ontving ik altijd snel uitgebreide en gedetailleerde feedback. Jullie hebben de papers naar een hoger niveau getild, ik had het niet zonder jullie gekund. Dank!

Dear co-authors: Marc Bardou, Isabel Portillo, Sarah Hoeck, Isabel de Brabander, Philippe Perrin, Catherine Dubois, Marloes Elferink, Valerie Lemmens and Marc Thelen. It was a pleasure to work with all of you; it gave me a broader perspective on colorectal cancer screening, inside and outside the Netherlands.

Lydia, Sten, Monique, Iris en alle andere RIVM betrokkenen, het was fijn samenwerken, bedankt! Het opleveren van de landelijke monitor en notities in combinatie met promoveren was ideaal. Jullie zijn de ultieme stok achter de deur. Jullie betrokkenheid, enthousiasme en bereikbaarheid zorgde voor een efficiënte samenwerking.

FSB en in het bijzonder Else-Mariëtte, in de eerste jaren van mijn promotie zag ik jullie wekelijks. Uren hebben we samen gezeten om de monitor en indicatoren goed in ScreenIT te krijgen. Else, jouw precisie heeft ervoor gezorgd dat alles goed verwerkt is. Bedankt voor het aanleveren van alle benodigde data voor onze landelijke monitor en evaluaties.

Arry, dank voor al jouw hulp wanneer er iets geregeld moest worden, vooral voor mijn verdediging. Jij loopt altijd een paar stappen voor ons uit.

Lieve kamergenootjes, in vier jaar tijd ben ik twee keer een verdieping naar boven verhuisd. Van een zespersoonskamer, naar een achtpersoonskamer, naar flexwerken. Het grote voordeel hiervan is dat ik veel collega's beter heb leren kennen. Daardoor kon ik altijd iemand vinden die mij kon helpen als ik ergens tegenaan liep of als ik zin in koffie had. Flexwerken is zo slecht nog niet.

Mijn GI-collega's: Frank, Miriam, Sonja, Steffie, Reinier, Maaike, Andrea, Amir, Dayna, Elleke, Carlijn, Arthur, Caroline, Lucie en Brechtje. It was great to be part of the biggest and smartest research group of MGZ. We're the best! Now I am working from home, I realise how important and relevant colleagues are. Thank you for all your help, support, and 'gezelligheid'.

Lieve Elleke and Dayna, our weekly gossip hour is amazing. It's great to share our experience about our long-distance relation with MGZ. *Elleke,* bedankt dat je de eerste versies van mijn introductie en discussie hebt gelezen, waardoor ik enorm.

Lieve Arthur, mijn kleine broertje. Het was een feest om met jou samen te werken, samen als de data-'sukkels'. Je bent een relaxed persoon en vindt alles mooi. Ik heb je nog nooit chagrijnig gezien, zelfs niet na het krijgen van negatieve feedback. Bedankt voor alle keren dat je dingen van mij hebt overgenomen, omdat ik weer eens zwanger was of ging verhuizen naar de andere kant van de wereld. Ik kijk uit naar jouw verdediging!

Lieve Carlijn, wat fijn dat jij bij MGZ kwam werken. Naast een gezellige, enthousiaste en slimme collega ben je vooral ook een lieve vriendin geworden. Naast lekker klagen over PhD-gerelateerde zaken, was het heerlijk om samen te trainen voor de triatlon, shoppen en zelfs chillen in Singapore. Het is meer dan verdiend dat jij vandaag naast mij staat te shinen!

Lieve vriendjes en vriendinnetjes, ik ga jullie niet bij naam noemen omdat ik bang ben dat ik iemand vergeet ☺. Maar jullie geven mijn leven kleur en zijn allemaal op jullie eigen manier belangrijk voor mij. Uit het oog is hopelijk niet uit het hart!

Lieve Janneke en Rosa, voor jullie een speciaal dankwoord. Jullie zijn die vriendinnen met wie ik hard kan lachen, maar die ook ongemakkelijke vragen stellen. Jullie maken mij tot een mooier mens, bedankt! Jans, jij bent vandaag de vertegenwoordiging van de kippetjes.

Lieve (schoon)familie: Oma Grietje, Sanne, Corstiaan, Tiemen, Jos, Liesbeth, Rogier, Marieke, Ilse, Niels, Jurre, Femke, en alle kleine neefjes en nichtjes. Ik ben blij met jullie allemaal. *Lieve Liesbeth,* bedankt voor alle uren oppassen, het was heerlijk werken als de mannetjes thuis met oma waren. *Lieve Femke,* bedankt voor het ontwerpen van de cover.

Lieve broer, zus en zusje. Ik voel me gelukkig omdat ik jullie zusje ben. Het is een fijne, gezellige en drukke bende met elkaar (die tijdens mijn promotie is uitgebreid met 7 kinderen ☺). Ik ben trots op jullie alle drie! *Lieve Roeland*, jij zorgde voor mijn afstudeerplek bij de MDL, wat indirect heeft geleid tot deze promotie plek. Door jouw voorbeeld wist ik dat promoveren vooral heel leuk is. Mooi om te zien hoe jij je werk als MDL-arts, je rol als man en vader en je uren wielrennen weet te combineren. *Lieve Mirjam*, onze hoogtepunten tijdens mijn promotie waren overduidelijk: twee keer tegelijk zwanger en een baby krijgen, fantastisch! Door jou was (en is) het helemaal niet zwaar om moeder te worden (en zijn). Je bent loyaal en hardwerkend, al jaren de beste fysio van Rijndam. *Lieve Judith*, je blijft altijd mijn kleine zusje, al ben je dat al lang niet meer. Zeker niet nu je ook moeder bent van mijn

kleinste neefje. Heerlijk om te zien hoe gepassioneerd je bent bij de dermatologie en hoe je daar van geniet. Ik hoop dat je papers ook eindigen in een mooi proefschrift!

Lieve pap en mam, bedankt voor jullie nooit aflatende aandacht en liefde. Jullie hebben mij laten zien dat ik geliefd ben om wie ik ben, niet om wat ik doe. Jullie waren denk ik net zo trots op mijn zwemdiploma als op mijn promotie. Jullie zijn een voorbeeld voor me, ik hoop dat ik net zo goed voor mijn eigen kinderen zal zorgen.

Lieve Jens en Faas, mijn twee kleine boeven. Jullie waren de mooiste hoogtepunten van mijn promotietraject. Ik heb altijd zin om te gaan werken, maar sinds jullie er zijn is thuiskomen een feest! Jullie werken ook relativerend, sinds jullie komst ben ik nooit meer gestrest. Ik ben dankbaar dat ik jullie moeder mag zijn, ik hou oneindig veel van jullie.

Lieve Harmen, de afgelopen jaren waren er meer life events dan publicaties; promoveren leek eigenlijk bijzaak. Van ons huurappartementje naar ons eigen (klus)huis in het pittoreske Overschie, geboorte Baby 1, geboorte Baby 2, en de emigratie naar Singapore. Het was een goede tijd, we zijn rijk gezegend. Ik hou van je! Dat er nog veel mooie herinneringen bij mogen komen.

Lieve God, wetenschappelijk niet te verklaren, maar voor mij dagelijkse realiteit. Leven met U zet alles in een ander perspectief. The best is yet to come!

ABOUT THE AUTHOR

Esther Zoutendijk was born in 1986 in Dirksland, the Netherlands. She completed her VWO at the secondary school Christelijke Scholengemeenschap Prins Maurits in Middelharnis. In 2005 she commenced her Nursing degree at Hogeschool Rotterdam. After obtaining her nursing qualification in 2009, she worked as a registered nurse at Maasstad Hospital in Rotterdam. In 2010 she moved to Perth, Australia, where she lived for two years. When Esther returned to the Netherlands in 2012 she undertook a two-year master programme in Health Sciences at the Vrije Universiteit in Amsterdam. Esther obtained her Master's degree (*cum laude)* in 2014. In February 2015 she started her PhD at the Department of Public Health at Erasmus MC University Medical Center focusing on the implementation of the Dutch colorectal cancer screening programme. During her PhD Esther also obtained her university teaching qualification (BKO). Esther currently lives in Singapore, together with her husband and two sons.

LIST OF PUBLICATIONS

This thesis

The second round of the Dutch colorectal cancer screening programme: impact of an increased FIT cut-off level.
Kooyker AI, Toes-Zoutendijk E, Opstal-van Winden AWJ, Spaander MCW, Buskermolen M, van Vuuren AJ, Kuipers EJ, van Kemenade FJ, Ramakers C, Thomeer MGJ, Dekker E, Nagtegaal ID, de Koning HJ, van Leerdam ME, Lansdorp-Vogelaar I.
Manuscript submitted.

Socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal screening programme with faecal immunochemical testing
van der Meulen MP, Toes-Zoutendijk E, Spaander MCW, Dekker E, Bonfrer JMG, van Vuuren AJ, Kuipers EJ, van Kemenade FJ, van Velthuysen MF, Thomeer MGJ, van Veldhuizen H, de Koning HJ, Lansdorp-Vogelaar I; Dutch National Colorectal Cancer Screening Working Group.
Manuscript submitted.

Adherence to screening in FIT-based colorectal cancer screening programmes in the NorthWest of Europe.
Toes-Zoutendijk E, Portillo I, Hoeck S, de Brabander I, Perrin P, Dubois C, van Leerdam ME, Lansdorp-Vogelaar I, Bardou M.
J Med Screen. 2019.

Incidence of interval colorectal cancer after negative result from first-round faecal immunochemical screening tests, by cutoff value and participant sex and age.
Toes-Zoutendijk E, Kooyker AI, Dekker E, Spaander MCW, Opstal-van Winden AWJ, Ramakers C, Buskermolen M, van Vuuren AJ, Kuipers EJ, van Kemenade FJ, Velthuysen MF, Thomeer MGJ, van Veldhuizen H, van Ballegooijen M, Nagtegaal ID, de Koning HJ, van Leerdam ME, Lansdorp-Vogelaar I; Dutch National Colorectal Cancer Screening Working Group.
Clin Gastroenterol and Hepatol. 2019.

Quality Monitoring of a FIT-Based Colorectal Cancer Screening Program.
Toes-Zoutendijk E, Bonfrer JMG, Ramakers C, Thelen M, Spaander MCW, Dekker E, van der Meulen MP, Buskermolen M, van Vuuren AJ, Kuipers EJ, van Kemenade FJ, van Velthuysen MF, Thomeer MGJ, van Veldhuizen H, van Ballegooijen M, de Koning HJ, van Leerdam ME, Lansdorp-Vogelaar I.
Clin Chem. 2019;65(3):419-426.

Stage distribution of screen-detected colorectal cancers in the Netherlands.
Toes-Zoutendijk E, Kooyker AI, Elferink MA, Spaander MCW, Dekker E, Koning HJ, Lemmens VE, van Leerdam ME, Lansdorp-Vogelaar I; LECO working group.
Gut. 2018;67(9):1745-1746.

Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels.
Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal ID, van der Meulen MP, van Vuuren AJ, Kuipers EJ, Bonfrer JMG, Biermann K, Thomeer MGJ, van Veldhuizen H, Kroep S, van Ballegooijen M, Meijer GA, de Koning HJ, Spaander MCW, Lansdorp-Vogelaar I;
Dutch National Colorectal Cancer Screening Working Group.
Gastroenterology. 2017 Mar;152(4):767-775.

Other

Modeling in Colorectal Cancer Screening: Assessing External and Predictive Validity of MISCAN-Colon Microsimulation Model Using NORCCAP Trial Results.
Buskermolen M, Gini A, Naber SK, Toes-Zoutendijk E, de Koning HJ, Lansdorp-Vogelaar I.
Med Decis Making. 2018;38(8):917-929.

National population screening for colorectal carcinoma in the Netherlands: results of the first years since the implementation in 2014.
Elferink MAG, Toes-Zoutendijk E, Vink GR, Lansdorp-Vogelaar I, Meijer GA, Dekker E, Lemmens VEPP.
Ned Tijdschr Geneeskd. 2018;162.

Colorectal cancer screening: Associations between information provision, attitudes and intended participation.
Brandhof SD, Fagerlin A, Hawley S, Toes-Zoutendijk E, Trevena L, McCaffery K, Korfage IJ.
Patient Educ Couns. 2018;101(3):546-550.

Bevolkingsonderzoek darmkanker in Nederland – Population screening for colon cancer in the Netherlands.
van Berkel AM, Lansdorp-Vogelaar I, Toes-Zoutendijk E, van Leerdam ME.
Ned Tijdschr Oncol. 2017;14:140-5.

Epidemiological trends among the population with chronic HCV infection in the Netherlands.
Maan R, Toes-Zoutendijk E, Veldt BJ, Hansen BE, van der Meer AJ, de Knegt RJ.
Antivir Ther. 2016;21(3):207-15.

PHD PORTFOLIO

Name PhD candidate:	Esther Toes-Zoutendijk
PhD period:	February 2015 - August 2019
Erasmus MC department:	Public Health
Promotor:	Prof. dr. Harry J. de Koning
Co-promotoren:	Dr. Iris Lansdorp-Vogelaar
	Dr. Monique E. van Leerdam

COURSES AND WORKSHOPS	Year	Workload (hours)
NIHES, Erasmus MC, Rotterdam		
Advanced topics in Decision Making in Medicine, Rotterdam	2015	40
Planning and evaluation of screening, Rotterdam	2015	40
Biostatistical Methods II: Classical Regression Models, Rotterdam	2015	120
Principles of research in medicine and epidemiology, Rotterdam	2016	8
Masterclass: Advances in Epidemiologic Analysis, Rotterdam	2016	4
Joint Models for Longitudinal and Survival Data, Rotterdam	2016	40
General courses, Erasmus MC, Rotterdam		
Wetenschappelijke Integriteit, Rotterdam	2016	12
Biomedical English Writing and communication, Rotterdam	2017	80
BKO	2017-2018	160

PRESENTATIONS		
Oral presentations		
United European Gastroenterology Week, Barcelona	2015	16
NVRO, Amersfoort	2015	16
Dutch Digestive Disease Days, Veldhoven	2017	16
World Endoscopy Organization: Colorectal Cancer Screening Committee, Barcelona	2017	16
Netwerk bijeenkomst bevolkingsonderzoek Noord, Groningen	2018	16
World Endoscopy Organization: Colorectal Cancer Screening Committee, Washington DC	2018	16
Interdisciplinary symposium 'innovations in oncology' - Berlin Summit, Berlin	2018	16
Dutch Digestive Disease Days, Veldhoven	2019	16
Poster Presentations		
Digestive Disease Week 2016, San Diego	2016	8
United European Gastroenterology Week 2017, Barcelona	2017	8
Digestive Disease Week 2018, Washington DC	2018	8

CONFERENCES AND MEETINGS	Year	Workload
International Cancer Screening Network Meeting 2015, Rotterdam	2015	24
Federadag 2015: Cancers and Numbers, Enschede	2015	8
World Endoscopy Organization: Colorectal Cancer Screening Committee, Barcelona	2015	8
United European Gastroenterology Week 2015, Barcelona	2015	24
Cancer Screening in the European Union - working group meeting (IARC), Lyon	2016	16
World Endoscopy Organization: Colorectal Cancer Screening Committee, San Diego	2016	8
Digestive Disease week 2016, San Diego	2016	40
NvvO Milestonedag: Preventie van kanker, Amsterdam	2016	8
World Endoscopy Organization: Colorectal Cancer Screening Committee, Vienna	2016	8
World Endoscopy Organization: Colorectal Cancer Screening Committee, Barcelona	2017	8
United European Gastroenterology Week 2017, Barcelona	2017	24
Congres Bevolkingsonderzoeken naar kanker, Utrecht	2017	8
World Endoscopy Organization: Colorectal Cancer Screening Committee, Washington	2018	8
Digestive Disease week 2018, Washington	2018	40

SEMINARS AND WORKSHOPS		
Presentatie training ARTESC, Rotterdam	2016	8
Research meetings department of Public Health Erasmus MC, Rotterdam	2015-2018	
Career Guidance Program, Rotterdam	09-2017 t/m 01-2018	80
Werkgroep FIT Kwaliteit, Capaciteit, Redactieraad, Monitoring & Evaluatie (KCMI), Utrecht	2015-2019	80

TEACHING	Year	Workload
Lecturing		
Planning and Evaluation of Screening, NIHES, Rotterdam	2017-2019	24
Supervising practicals and tutoring		
Community project, Bachelor geneeskunde jaar 3, Rotterdam	2016	16
Primaire preventie in de artsenpraktijk, Bachelor geneeskunde jaar 3, Rotterdam	2016	24
Academisch redeneren en schrijven I&II, Bachelor geneeskunde jaar 2, Rotterdam	2016	32
Primaire preventie in de artsenpraktijk, Bachelor geneeskunde jaar 3, Rotterdam	2017	8
Bachelor essays, Bachelor geneeskunde jaar 3, Rotterdam	2017	28
Keuzen in de zorg, Bachelor geneeskunde jaar 3, Rotterdam	2017	16
Ontwikkelen e-module geneeskunde onderwijs 'screening', Rotterdam	2018	24
NIHES 'Planning and Evaluation of Screening' vernieuwen en ontwikkelen, Rotterdam	2018	40
GRANT AND REVIEW ACTIVITIES		
Grant proposal ZonMW	2017	32
Grant proposal Smarter Choices Better Health Erasmus University	2018	32
Grant proposal MLDS	2019	32
Reviewer for Eur J of Epi, BMC Medicine, GUT, Clin Gastroenterol, JAMA, Journal of Gastro and Hep	2017-2019	30
Reviewer of grant proposal (Dutch subsidy provider)	2018	16