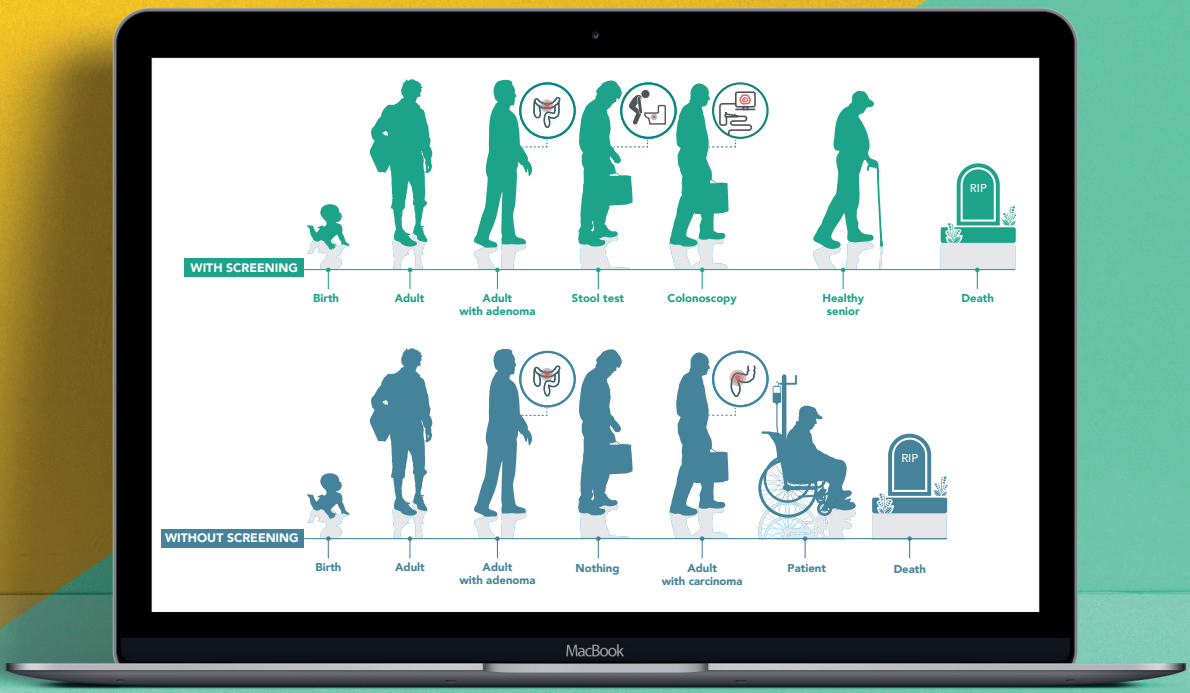
IMPROVING THE PERFORMANCE AND EFFICIENCY OF THE COLORECTAL CANCER SCREENING PROGRAMME IN THE NETHERLANDS



Miriam van der Meulen

Improving the performance and efficiency of the colorectal cancer screening programme in the Netherlands

Miriam Petra van der Meulen

ISBN: 978-94-028-2055-3

Printed by Ipskamp Printing, proefschriften.net
Layout and design: Lara Leijtens, persoonlijkproefschrift.nl
Cover design: Steven Olde Hengel

Improving the Performance and Efficiency of the Colorectal Cancer Screening Programme in the Netherlands

Verbeteren van het presteren en de efficiëntie van het bevolkingsonderzoek dikke darmkanker in Nederland.

Proefschrift

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

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

en volgens het besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

woensdag 24 juni 2020 om 11.30 uur

Miriam van der Meulen

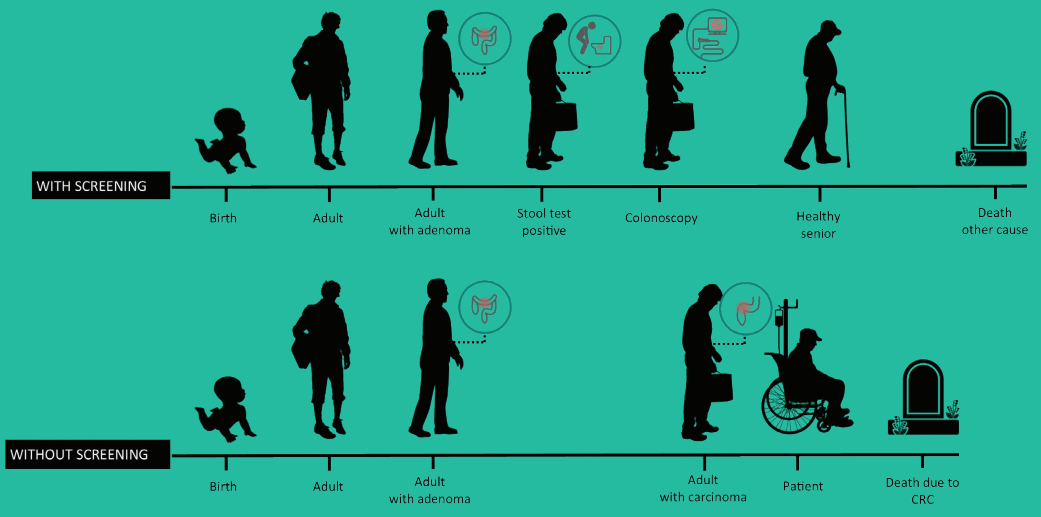
geboren op woensdag 1 februari 1984
te Leeuwarden

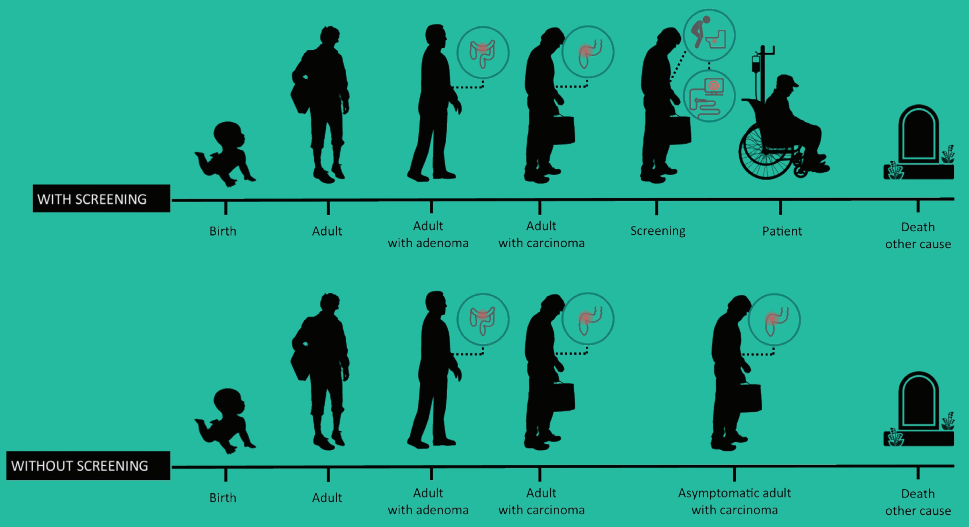
Erasmus University Rotterdam

Promotiecommissie:

Promotor:	Prof. dr. H.J. de Koning
Overige leden:	Prof dr. E. Dekker Prof. dr. A.E. Kunst Dr. M.C.W. Spaander
Co-promotor:	Dr. I. Lansdorp - Vogelaar

CONTENTS


WITH SCREENING
Birth
Adult
Adult with adenoma
Stool test positive
Colonoscopy
Healthy senior
Death other cause
WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Patient
Death due to CRC


WITH SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Screening
Patient
Death other cause
WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Asymptomatic adult with carcinoma
Death other cause

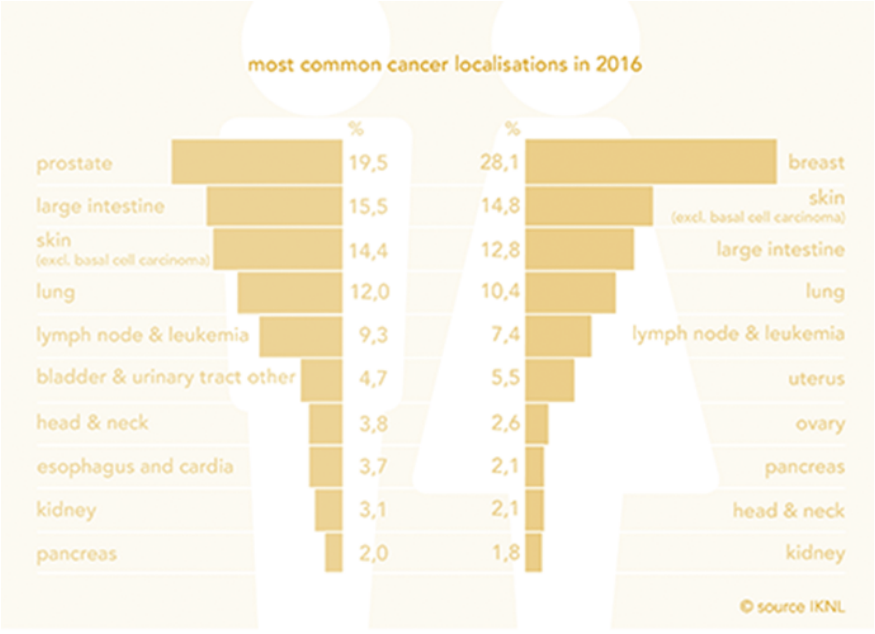
Chapter 1

General introduction

EPIDEMIOLOGY OF COLORECTAL CANCER

Colorectal cancer (CRC) is the second-most common cause of cancer mortality in the western world. The Netherlands had the fifth highest incidence rate of the world in 2018.[1] In the Netherlands, colorectal cancer is the second most common cancer in men, after prostate cancer, and the third most common cancer in women, only breast and skin cancer are more common than CRC (Figure 1).[2] In absolute numbers, the incidence has been steadily rising from 7,100 cases in 1990 to 13,028 cases in 2013 (the year before the introduction of screening) (Figure 2), while he European age-standardized rate (ESR) increased from 45 to 56 cases per 100,000 individuals. The highest absolute number of CRC's are diagnosed in people between 65 and 79 years old, but the highest incidence rate of CRC per 100,000 in the population occurs between 80 and 84 years old (Figure 3).

Figure 1: **The most common cancer localisation in the Netherlands in 2016**



The number of fatalities due to CRC also increased from 3,900 deaths in 1990 to 5,000 deaths in 2013 (Figure 4), albeit the increase was relatively smaller than the increase in incidence due to improved survival. The 5-year survival of colorectal cancer is 59% and depends strongly on the stage at diagnosis. Five-year survival for stage I is 92%, while 5-year survival for stage IV is only 9%. Currently, almost half of the patients have lymph node or distant metastasis at the time of diagnosis (stage III or IV).

Figure 2: **Number of CRC diagnosis in the Netherlands from 1989 to 2015.**

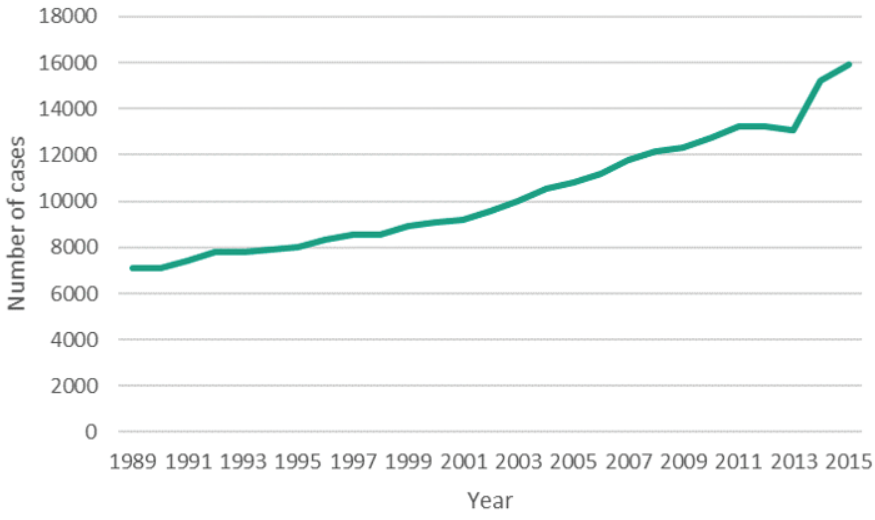
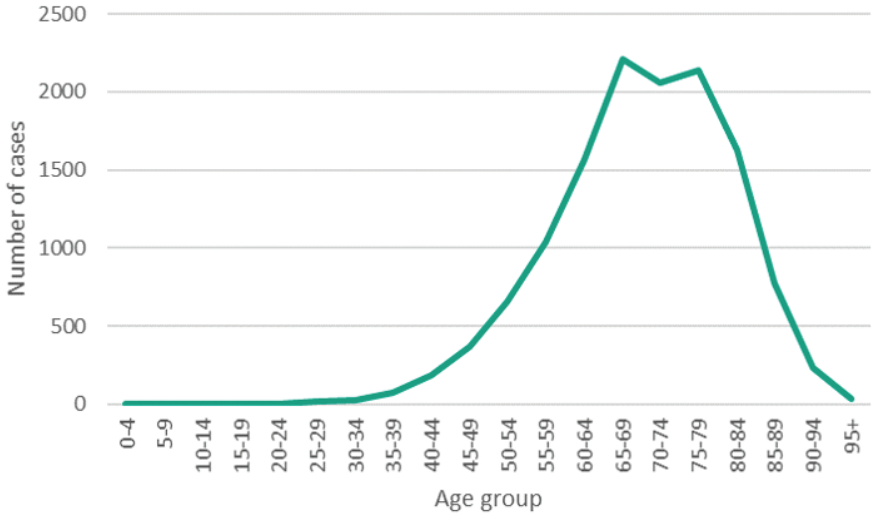


Figure 3: **Number of CRC diagnosis per 10,000 individuals in different age groups in the Netherlands in 2013.**

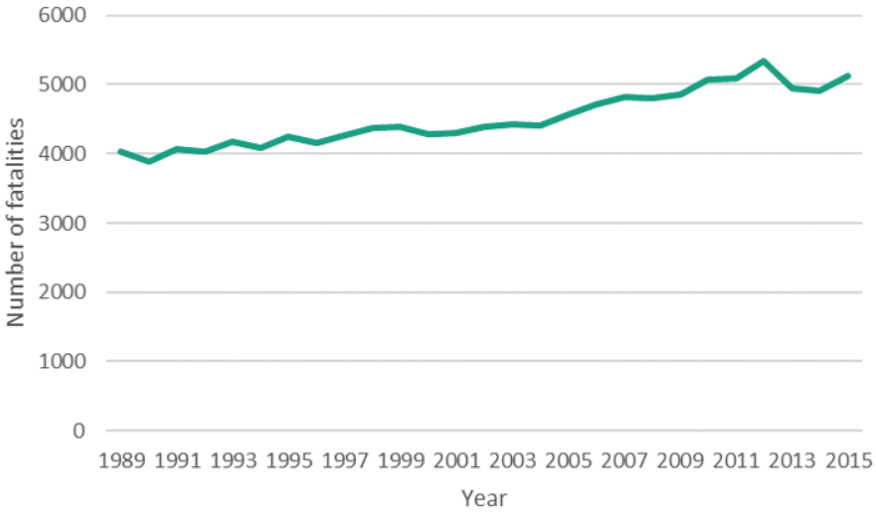


Men have a higher incidence of CRC than women. In 2013, 7335 men and 5693 women were diagnosed with CRC. The cumulative lifetime incidence of CRC for men is 7.49% versus 6.24% for women. The survival does not differ between men and women.

The incidence of CRC does not differ for different levels of socioeconomic status (SES) in the Netherlands.[3] The mortality due to CRC, however, does differ: individuals

with a lower SES (more deprivation, lower income, lower level of education) have a higher probability to ever die from CRC.

Figure 4: **Number of fatalities due to CRC in the Netherlands from 1989 to 2015.**



NATURAL HISTORY OF COLORECTAL CANCER

The majority of colorectal cancers is believed to develop from a precursor lesion, the adenoma.[4] About 30% of adults between the age of 50 and 75 years old has adenomas in their colorectum. Only a small percentage of these adenomas will eventually develop into a CRC.

If an adenoma has developed into CRC, it mostly does not give symptoms right away and is called preclinical CRC. If symptoms are present, it will last on average 3-5 months before CRC is diagnosed and treated due to patients, doctors and treatment delay.[5]

Recently, an alternative pathway to CRC is described, called the sessile serrated polyp pathway.[6] Sessile serrated polyps are often flat or sessile, may be covered with mucus and are more prevalent in the proximal and rectosigmoid colon. Therefore, the sensitivity to detect these lesions might be lower for several screening tests. Also, differences in the progression from adenomas to cancer may exists between traditional adenomas and sessile serrated polyps, but evidence on the pathway to cancer is still scarce.

PREVENTION TO REDUCE MORBIDITY AND MORTALITY FROM CRC

Three types of prevention exist to limit the morbidity and mortality of CRC: primary, secondary and tertiary prevention. In short, primary prevention aims to prevent the disease to develop by reducing or eliminating etiological factors of the disease and consists mostly of lifestyle interventions. Secondary prevention aims to detect (precursors of) disease in early stages to prevent worsening of or mortality due to the disease. Tertiary prevention aims to prevent or restrict the consequences of already diagnosed disease. The three types of prevention for CRC are discussed in short below. This thesis focuses on secondary prevention.

Primary prevention

Known risk factors for developing colorectal cancer are dietary factors, smoking, alcohol use, lack of physical activity and obesity.[7-10] Together, these risk factors could be a large contributor to the development of colorectal cancer. For instance, smoking is on average (over different countries) responsible for approximately 7% of CRC (population attributable fraction), obesity for 12% of CRC, while alcohol use is responsible for 15% and 4% of CRCs in males and females, respectively.[10] Lifestyle interventions targeting these risk factors could therefore reduce the risk of CRC.

Secondary prevention

Screening

With screening, if the precursor lesion is removed, colorectal cancer can be prevented. If CRC is not prevented by screening, it can be detected in an earlier stage. Because treatment will then also take place at an earlier stage, it can be easier to resect the tumor and there is a higher probability of complete resection and a lower probability of side effects. Also, there is a lower probability of lymphatic or distant metastasis., Thus, prevention and early detection of CRC can improve the survival and limit the need for harmful treatments.

In 1968, Wilson and Junger described criteria for a screening programme, which are shown in Figure 5 (Box 1). Over time, new criteria have been added to these “classic” criteria (Box 2). An important new screening criterium is that there should be scientific evidence of screening effectiveness. Several randomized controlled trials have demonstrated a CRC mortality reduction by screening ranging from 15%-33% with two different screening tests.[11-15] Another screening criterium is that the overall benefits of screening should outweigh the harm. Potential harms are anxiety due to a false-positive test, burden or complications due to a screening test, or its follow-up, and overdiagnosis. The burden and complications of screening are dependent on the choice of screening test and discussed below. Overdiagnosis is the diagnosis of disease

that would not have caused symptoms or problems in the absence of screening. If early diagnosis does not prolong life, the diagnosis could be harmful due to unnecessary treatment and the knowledge of having fatal disease.

Surveillance

Individuals with adenomas are at increased risk of developing metachronous adenomas and CRC, even after the adenomas have been completely removed.[16-18] Therefore, surveillance with colonoscopy is recommended for individuals that have had a polypectomy.[6, 19] The frequency of surveillance and the compliance to surveillance recommendations are important, since too little surveillance has the risk of decreasing the preventive effect of colonoscopy for CRC, while too intensive surveillance exposes the patient to unnecessary risks and burden and wastes colonoscopy and financial resources. However, research shows that surveillance is currently often not used efficiently. Several surveys and real-life data showed suboptimal compliance to guidelines.[20-22] In the Netherlands, the compliance to the guideline was also reported to be low.[21] In that survey, gastroenterologists indicated to deviate from the guideline because the surveillance interval was based on number of adenomas only, while in literature, additional risk factors for metachronous lesions were found. To accommodate clinical expertise and clinical evidence on the development of metachronous lesions, a new surveillance guideline was therefore introduced in 2013 in the Netherlands, incorporating size and location and histology of adenomas as well as presence of large serrated lesions.[23] Through a score chart these adenoma characteristics are combined into a risk score (0 - 5) to optimize the risk stratification of patients and recommend a 3 or 5-year interval or no surveillance for individuals with only one low-risk adenoma.

Tertiary prevention

The possibilities for management of CRC depends on various factors. The most important factor is the stage of a CRC. The stage has three components, the (size of the) primary tumor (T), status of the regional lymph nodes (N), and distant metastasis (M), which together are combined in stage groups from I to IV. The only curative option is surgery, which is only possible if the tumor is localized to the tumor wall or regional lymph nodes (approximately 80% of the cases). To restore the function of the bowel, an anastomosis is usually used, while in some cases, a colostomy or ileostomy is necessary. Only if the resection margins are possibly compromised neoadjuvant chemotherapy has a place in the management of CRC.[24, 25] Patients that already underwent potentially curative surgery could benefit from postoperative chemotherapy if they have stage III disease, the choice of postoperative chemotherapy for stage II patients is controversial. After treatment, posttreatment surveillance takes place. For stage I, this consists only of

periodic consultations and colonoscopy. For stage II and III, this is extended with serial assay of the serum concentrations of the tumor marker carcinoembryonic antigen (CEA) and annual surveillance computed tomography (CT) scans.

If the tumor is already metastasized at presentation (approximately 20%), prognosis is poor. Even with the major advances in systemic chemotherapy, fewer than 20 percent is still alive at five years.[26]

Figure 5: **Criteria for screening**

Box 1. Wilson and Jungner classic screening criteria[1]

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a "once and for all" project.

Box 2. Synthesis of emerging screening criteria proposed over the past 40 years

- The screening programme should respond to a recognized need.
- The objectives of screening should be defined at the outset.
- There should be a defined target population.
- There should be scientific evidence of screening programme effectiveness.
- The programme should integrate education, testing, clinical services and programme management.
- There should be quality assurance, with mechanisms to minimize potential risks of screening.
- The programme should ensure informed choice, confidentiality and respect for autonomy.
- The programme should promote equity and access to screening for the entire target population.
- Programme evaluation should be planned from the outset.
- The overall benefits of screening should outweigh the harm.

SCREENING PROGRAM

Screening test

Several different tests are suitable for use in colorectal cancer screening. Not one of these tests is considered the most optimal test, because the comparison of the tests differs for different outcome parameters (eg yield of a screening round or long-term cost-effectiveness) and is dependent of country and setting specific participation. Therefore, several different tests are currently in use for colorectal cancer screening.

Colonoscopy is considered the reference standard for diagnosis of adenomas and CRC and in addition, it allows for direct removal of colonic lesions. Therefore, if a participant tests positive on a screening test other than colonoscopy, that participant still has to undergo colonoscopy. Its use as a primary screening test has drawbacks: the test can be burdensome because a bowel preparation is required and this can result in relatively low participation. It is also costly and a large endoscopy capacity is needed.

These drawbacks can be reduced by selecting individuals with increased risk to undergo colonoscopy, for instance by incorporating a less invasive primary test and offering colonoscopy only to test positives. The faecal occult blood test (FOBT) is currently the most often used screening test for CRC screening, other test modalities that are used, besides colonoscopy, are sigmoidoscopy and computed tomography colonography (CTC).[27]

FOBT is a stool-based test to determine the presence of blood in the stool. An FOBT test can be mailed and performed at home. A small sample of stool should be collected by the participant in a special (small) container and send back to the screening organization or doctor. The faecal immunochemical test (FIT) is a variant of the FOBT that can generate a quantitative result. Therefore, the cut-off for referral to colonoscopy can be adjusted to a preferred level. Because the test can be performed at home and is non-invasive, the test is convenient for a participant to do. However, if the test is positive, a colonoscopy should still be done. Another disadvantage is the suboptimal sensitivity of FOBT especially for adenomas, which can be compensated by repeated testing.

CTC involves obtaining multiple, thin-slice CT data and uses computers to construct images of the colon in two and three dimensions. A bowel preparation is still required for CTC and CTC is also burdensome due to exposure to radiation and inflation of the colon. CTC may yield incidental radiologic findings in other organs that may require additional testing. It is unclear if the resulting procedures of the detection of such incidental findings improves health outcomes or results in additional harms.

Sigmoidoscopy is an endoscopic procedure just like colonoscopy, but a sigmoidoscope reaches from the rectum until the splenic flexure instead of the

ileum with colonoscopy. Thus, sigmoidoscopy examines only the distal portion of the colon, while 41 to 45 percent of CRCs are in the right side of the colon. Therefore, sigmoidoscopy may miss some of the polyps.[28] Sigmoidoscopy has similar disadvantages as colonoscopy. However, because the bowel preparation is less intensive than for colonoscopy or CTC, and sedation is not needed, the burden for the patient is lower. If a sigmoidoscopy is positive, a colonoscopy will be advised for further evaluation of the complete colon. Already during sigmoidoscopy, small polyps can be removed, however, removal of larger polyps (>1.0 cm) is preferred to be done during a subsequent colonoscopy.

Type of screening programme

Colorectal cancer screening is widely adopted across the world; however, the screening programs differ in the way they are organized, the choice of screening test and the age range and interval of screening.

Screening can be either opportunistic or centrally organised. In organised screening programmes, large numbers of people are actively invited to take part in screening and everyone who takes part is offered the same services, information and support. Opportunistic screening happens when screening is recommended or reimbursed in general, but it is left up to the individual to organize their screening. That is: the initiative for screening lies with the individuals to ask their doctor or health professional for a check or test. Unlike an organised screening programme, opportunistic screening programmes often lack quality assurance, monitoring and evaluation. About half of the countries with colorectal cancer screening have an organised screening programme, while others, such as the US and Germany have opportunistic colorectal cancer screening.[27]

Dutch screening programme

In the Netherlands, a population-based colorectal cancer screening programme was introduced in 2014. Before deciding to introduce this screening programme, pilot-studies were performed to investigate the acceptation and performance of different screening tests in the Dutch population. In 2008, a randomized controlled trial started comparing the yield and participation to sigmoidoscopy, gFOBT and FIT. FIT was observed to have the highest participation rate and consequently also the highest yield per invitee.[29] Therefore, several other studies involving FIT were performed, including a comparison of 1- and 2 sample FIT screening and the effect of different screening intervals with FIT in a subsequent round.[30, 31] In addition, a randomized controlled trial compared a single round of colonoscopy screening to a single round of CTC screening.[32] These pilot-studies showed FIT is most acceptable to the Dutch population with a participation rate of

up to 60%-62%, compared with 47%-50% for gFOBT, 32% for sigmoidoscopy, 22% for colonoscopy and 34% for CTC. Modeling studies with MISCAN-Colon based on the FIT pilots showed it is most cost-effective to screen with 1-sample FIT with a low cut-off.[33] If colonoscopy capacity is too limited for the preferred screening strategy, the most cost-effective adaptation of the screening strategy is to raise the FIT cut-off.[34]

Based on these findings and on the preferred balance between true and false positives, biennial FIT was introduced in the Netherlands in 2014 with a gradual roll-out period of five years. The target population consists of individuals aged 55 to 75 years and consist of approximately 2 million invitees yearly. 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). After 42 days a reminder is sent automatically to non-responders. To optimize the balance between true-positive and false-positive test results, initially a cut-off of 15 ug/g faeces was chosen. During the first year of screening, the cut-off was increased to 47 ug/g faeces, because of higher than expected positivity rate, a lower than expected PPV, and limited colonoscopy capacity.[35]

If the FIT result is equal or greater than the cut-off level, the participant is invited for a pre-colonoscopy intake interview in an accredited colonoscopy center nearby. Participants whose sample is unreliable or not assessable are sent a new test. Colonoscopy is the standard diagnostic follow-up test. All colonoscopies are performed by accredited endoscopists who perform at least 300 colonoscopies each year. All detected polyps are to be removed and sent for pathologic review. In case of advanced adenoma (AA) or CRC, the participant is referred for further treatment and surveillance.

All data are collected in a national information system of the CRC screening programme (ScreenIT). ScreenIT includes personal details (like sex, 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.

OUTLINE OF THE THESIS AND RESEARCH QUESTIONS

In this thesis, three chapters use data from the Dutch pilot studies in MISCAN-Colon to optimise colorectal cancer screening in the Netherlands. In one chapter, we use data from the national screening program to explore SES differences in screening participation and yield. In the last chapter, we investigate the compliance of gastroenterologists to the guideline of surveillance after polypectomy.

Research questions:

Chapter 2) Are there differences in FIT performance between men and women?

Chapter 3) Do men and women need to be screened differently with fecal immunochemical testing from a cost-effectiveness perspective?

Chapter 4) Do systematic false-negative fecal immunochemical test results exist and what are their implications for screening effectiveness?

Chapter 5) What is the comparative cost-effectiveness of CTC versus colonoscopy screening with assumed data on attendance and costs from a randomized controlled screening trial in a dedicated screening setting?

Chapter 6) What are the socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal screening programme with faecal immunochemical testing.

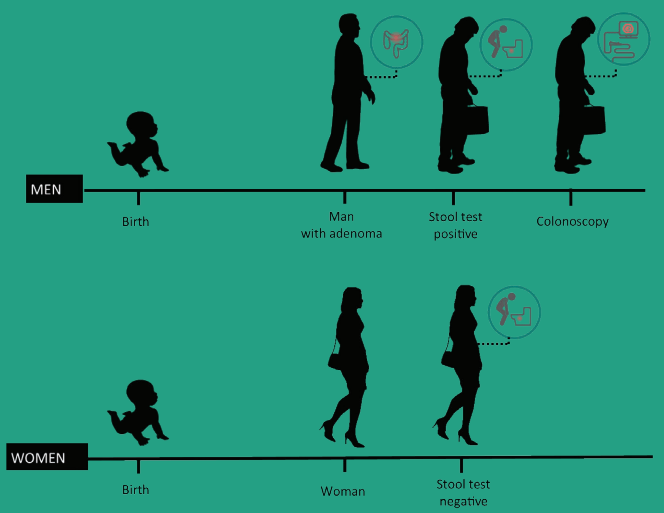
Chapter 7) How do gastroenterologists interpret and comply to the updated risk-stratified guideline for surveillance after polypectomy?

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
2. Comprehensive Cancer Centre: Dutch cancer registry; [Available from: http://cijfersoverkanker.nl/.
3. Aarts MJ, Lemmens VE, Louwman MW, Kunst AE, Coebergh JW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. European journal of cancer (Oxford, England : 1990). 2010;46(15):2681-95.
4. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50(1):113-30.
5. Weller D, Menon U, Zalounina Falborg A, Jensen H, Barisic A, Knudsen AK, Bergin RJ, Brewster DH, Cairnduff V, Gavin AT, Grunfeld E, Harland E, Lambe M, Law RJ, Lin Y, Malmberg M, Turner D, Neal RD, White V, Harrison S, Reguilon I, Vedsted P. Diagnostic routes and time intervals for patients with colorectal cancer in 10 international jurisdictions; findings from a cross-sectional study from the International Cancer Benchmarking Partnership (ICBP). BMJ open. 2018;8(11):e023870.
6. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology. 2012;143(3):844-57.
7. Lanting CI, de Vroome EM, Elias SG, van den Brandt PA, van Leeuwen FE, Kampman E, Kiemeney LA, Peeters PH, de Vries E, Bausch-Goldbohm RA. [Contribution of lifestyle factors to cancer: secondary analysis of Dutch data over 2010 and a projection for 2020]. Nederlands tijdschrift voor geneeskunde. 2014;159:A8085.
8. Lee J, Jeon JY, Meyerhardt JA. Diet and lifestyle in survivors of colorectal cancer. Hematology/oncology clinics of North America. 2015;29(1):1-27.
9. Richardson A, Hayes J, Frampton C, Potter J. Modifiable lifestyle factors that could reduce the incidence of colorectal cancer in New Zealand. The New Zealand medical journal. 2016;129(1447):13-20.
10. Whiteman DC, Wilson LF. The fractions of cancer attributable to modifiable factors: A global review. Cancer epidemiology. 2016;44:203-21.
11. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. Cancer. 1986;58(2):397-403.
12. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. The American journal of gastroenterology. 2008;103(6):1541-9.
13. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet (London, England). 1996;348(9040):1467-71.
14. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. Journal of the National Cancer Institute. 1999;91(5):434-7.
15. Shroff J, Thosani N, Batra S, Singh H, Guha S. Reduced incidence and mortality from colorectal cancer with flexible-sigmoidoscopy screening: a meta-analysis. World journal of gastroenterology. 2014;20(48):18466-76.

16. Cottet V, Jooste V, Fournel I, Bouvier AM, Faivre J, Bonithon-Kopp C. Long-term risk of colorectal cancer after adenoma removal: a population-based cohort study. Gut. 2012;61(8):1180-6.
17. Leung K, Pinsky P, Laiyemo AO, Lanza E, Schatzkin A, Schoen RE. Ongoing colorectal cancer risk despite surveillance colonoscopy: the Polyp Prevention Trial Continued Follow-up Study. Gastrointestinal endoscopy. 2010;71(1):111-7.
18. Martinez ME, Baron JA, Lieberman DA, Schatzkin A, Lanza E, Winawer SJ, Zauber AG, Jiang R, Ahnen DJ, Bond JH, Church TR, Robertson DJ, Smith-Warner SA, Jacobs ET, Alberts DS, Greenberg ER. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. Gastroenterology. 2009;136(3):832-41.
19. Atkin WS, Valori R, Kuipers EJ, Hoff G, Senore C, Segnan N, Jover R, Schmiegel W, Lambert R, Pox C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. Endoscopy. 2012;44 Suppl 3:Se151-63.
20. Huppertz J, Coriat R, Leblanc S, Gaudric M, Brezault C, Grandjouan S, Chaput U, Prat F, Chaussade S. Application of ANAES guidelines for colonoscopy in France: a practical survey. Gastroenterologie clinique et biologique. 2010;34(10):541-8.
21. Mulder SA, Ouwendijk RJ, van Leerdam ME, Nagengast FM, Kuipers EJ. A nationwide survey evaluating adherence to guidelines for follow-up after polypectomy or treatment for colorectal cancer. Journal of clinical gastroenterology. 2008;42(5):487-92.
22. van Kooten H, de Jonge V, Schreuders E, Sint Nicolaas J, van Leerdam ME, Kuipers EJ, Veldhuyzen van Zanten SJ. Awareness of postpolypectomy surveillance guidelines: a nationwide survey of colonoscopists in Canada. Canadian journal of gastroenterology = Journal canadien de gastroenterologie. 2012;26(2):79-84.
23. NVMDL. CBO guideline [Available from: http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf.
24. (NCCN) NCCN. NCCN clinical practice guidelines in oncology [Available from: https://www.nccn.org/professionals/physician_gls.
25. Specialisten FM. Richtlijn Colorectaal carcinoom (CRC) 2019 [Available from: https://www.oncoline.nl/colorectaalcarcinoom.
26. Heinemann V, von Weikersthal LF, Decker T, Kiani A, Vehling-Kaiser U, Al-Batran SE, Heintges T, Lerchenmuller C, Kahl C, Seipelt G, Kullmann F, Stauch M, Scheithauer W, Hielscher J, Scholz M, Muller S, Link H, Niederle N, Rost A, Hoffkes HG, Moehler M, Lindig RU, Modest DP, Rossius L, Kirchner T, Jung A, Stintzing S. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. The Lancet Oncology. 2014;15(10):1065-75.
27. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, Kuipers EJ. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49.
28. Institute NC. SEER*Stat Database [Available from: https://seer.cancer.gov/data/.
29. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JC, van der Togt AC, Habbema JD, Kuipers EJ. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59(1):62-8.
30. van Roon AH, Goede SL, van Ballegooijen M, van Vuuren AJ, Looman CW, Biermann K, Reijerink JC, Mannetje H, van der Togt AC, Habbema JD, van Leerdam ME, Kuipers EJ. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut. 2013;62(3):409-15.

31. van Roon AH, Wilschut JA, Hol L, van Ballegooijen M, Reijerink JC, t Mannetje H, Kranenburg LJ, Biermann K, van Vuuren AJ, Francke J, van der Togt AC, Habbema DJ, van Leerdam ME, Kuipers EJ. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2011;9(4):333-9.
32. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, van de Vijver MJ, Biermann K, Thomeer M, van Leerdam ME, Fockens P, Stoker J, Kuipers EJ, Dekker E. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. The Lancet Oncology. 2012;13(1):55-64.
33. Wilschut JA, Hol L, Dekker E, Jansen JB, Van Leerdam ME, Lansdorp-Vogelaar I, Kuipers EJ, Habbema JD, Van Ballegooijen M. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology. 2011;141(5):1648-55.e1.
34. Wilschut JA, Habbema JD, van Leerdam ME, Hol L, Lansdorp-Vogelaar I, Kuipers EJ, van Ballegooijen M. Fecal occult blood testing when colonoscopy capacity is limited. Journal of the National Cancer Institute. 2011;103(23):1741-51.
35. Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I, van der Meulen MP, van Vuuren AJ, Kuipers EJ, Bonfrer JM, Biermann K, Thomeer MG, van Veldhuizen H, Kroep S, van Ballegooijen M, Meijer GA, de Koning HJ, Spaander MC, Lansdorp-Vogelaar I. 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.


MEN
Birth
Man
with adenoma
Stool test
positive
Colonoscopy
WOMEN
Birth
Woman
Stool test
negative

Chapter 2

Gender differences in Fecal immunochemical test performance for early detection of colorectal neoplasia.

Atija Kapidzic, Miriam P. van der Meulen, Lieke Hol, Aafke H.C. van Roon, Caspar W.N. Looman, Iris Lansdorp-Vogelaar, Marjolein van Ballegooijen, Anneke J. van Vuuren, Jacqueline C.I.Y. Reijerink, Monique E. van Leerdam, and Ernst J. Kuipers

Clin Gastroenterol Hepatol. 2015 Aug;13(8):1464-71.e4. doi: 10.1016/j.cgh.2015.02.023. Epub 2015 Feb 24.

ABSTRACT

Background & Aims:

Fecal immunochemical tests (FITs) are widely used in colorectal cancer screening. Programs use the same fecal hemoglobin threshold for colonoscopy referral for men and women, but it is unclear whether FIT performs equally in both sexes. We therefore assessed FIT performance in men and women.

Methods:

A prospective cohort study was performed, in which in total 10.008 average-risk subjects (aged 50-74 years) were invited for first and 8.316 average-risk subjects (aged 51-74 years) for second round screening with a single FIT. Subjects with a hemoglobin (Hb) level of ≥10 µg Hb/g (or ≥50 ng/ml) feces were referred for colonoscopy. The test characteristics were assessed by sex for a range of FIT cut-offs.

Results:

In total 59.8% of men and 64.6% of women participated in the first round ($p<0.001$). At a cut-off level of 10 µg Hb/g, the positivity rate was significantly higher among men (10.7%) compared to women (6.3%, $p<0.001$) in the first round. The detection rate of advanced neoplasia was 4.4% for men and 2.2% for women ($p<0.001$) in the first round. The positive predictive value for advanced neoplasia in the first round was 42% for men and 37% for women ($p=0.265$). A significantly higher false-positive rate (FPR) in men (6.3%) than in women (4.1%, $p<0.001$) was found. Similar differences in these test characteristics were seen in the second round.

Conclusions:

At a cut-off level of 10 µg Hb/g the FIT positivity rate was higher in men, reflected by both a higher detection rate and a higher FPR. The use of the same cut-off value in men and women in FIT screening is recommended based on equal test performance in terms of positive predictive value.

INTRODUCTION

Screening by means of a guaiac-based fecal occult blood test (FOBT) reduces colorectal cancer (CRC)-related mortality [1]. More recently, fecal immunochemical tests (FIT) proved more effective than guaiac-based FOBT due to both a higher uptake and higher detection rate of advanced neoplasia [2-4]. This explains the strong worldwide interest in fecal immunochemical tests as a primary screening tool [5-9]. Until now, similar FIT screening regimens are applied in men and women despite eminent sex disparities in prevalence and anatomic distribution of advanced neoplasia.

Several colonoscopy-based screening studies have reported a higher incidence and prevalence of advanced neoplasia in men compared to women [10-13]. The positive predictive value and detection rate of both FOBTs depend on the prevalence of advanced neoplasia in the tested population. As a consequence, guaiac-based FOBT screening results in a lower positivity and detection rate and may result in a higher proportion of false positive test results in women [14-16].

A Scottish gFOBT screening study reported more screen-detected CRCs in men (64.5%) compared to women (35.5%), whereas the number of interval CRCs was similar in both groups (men: 49.8% vs. women: 50.2%) [17]. These data suggest that gFOBT is less sensitive when used in women. This finding was confirmed by a German study, where subjects received a FOBT (gFOBT or FIT) prior to a screening colonoscopy. The authors found a substantial higher sensitivity and positive predictive value in men than in women for both FOBTs [18]. Another study which compared FIT with primary screening colonoscopy also found a higher sensitivity of FIT in men [19]. Aforementioned data were obtained from studies with colonoscopy as a primary screening tool and might have a different underlying risk than the (screening-naïve) population screened with FIT.

Data on gender differences in a population-based setting with FIT as a primary screening tool are lacking. In this study we therefore determined potential gender differences in performance of FIT in an average risk, screening-naïve Dutch population.

MATERIALS AND METHODS

This study was based on the CORERO-I and –II studies, the primary results of which have been described elsewhere [4, 20]. In brief, 10.008 (aged 50-74 years) were approached for first and 8.316 screenees (aged 51-74 years) for second round screening. The demographic data of all invitees were obtained from municipal population registers in the wider Rotterdam region. Random samples were taken from the target population by a computer-generated algorithm (Tenalea, Amsterdam, the Netherlands). Since there was no CRC screening program at the time of the trial in the Netherlands, the target

population was screening-naïve when first approached. Individuals with a history of inflammatory bowel disease or CRC, as well as those who had undergone a colonoscopy, sigmoidoscopy or barium contrast enema in the last 3 years and those with an estimated life expectancy of less than 5 years were excluded from the study. Subjects were not invited for the second screening round in case of a positive FIT in the first screening round, when they had become older than 74 years of age, when they had moved out of the region, or when they had died. Recruitment took place between November 2006 and December 2010.

Interventions

With each screening round, one FIT (OC-Sensor Micro, Eiken Chemical Co, Tokyo, Japan) was sent by mail to collect a single sample of one bowel movement, after which it was returned by mail. The test was considered positive when the hemoglobin concentration in the FIT sample was 10 µg Hb/g feces, which corresponds to 50 ng/ml. Subsequently, a hemoglobin concentration of 40 µg Hb/g feces corresponds with 200 mg/ml. If the sample return time (i.e. the interval in days between fecal sampling and FIT laboratory delivery) was longer than 7 days and there was a positive FIT, the participant was referred for colonoscopy. If the sample return time was longer than 7 days and there was a negative FIT, the participant received a new test, because hemoglobin could have degraded during this long return time. In case of a positive test, the participant was referred for colonoscopy (Appendix 1). In the second round, study subjects were divided over three groups to undergo repeated FIT testing at different screening intervals (i.e. one, two and three years, respectively) [20]. Based on these results, a two-year interval was applied to all groups in the third screening round.

Follow-up evaluation

Subjects with a positive FIT were scheduled for colonoscopy within 4 weeks. All colonoscopies were performed by experienced endoscopists, who had performed over 1000 colonoscopies. The maximum reach of the endoscope, adequacy of bowel preparation as well as the characteristics and location of any polyps were recorded. Experienced gastrointestinal pathologists evaluated all removed polyps according to the World Health Organization Classification of Tumors [21]. Patients with a positive colonoscopy entered a surveillance program according to guidelines of the Dutch Society of Gastroenterology, while subjects with a negative colonoscopy were referred back to the screening program, but were considered not to require FIT screening for ten years.

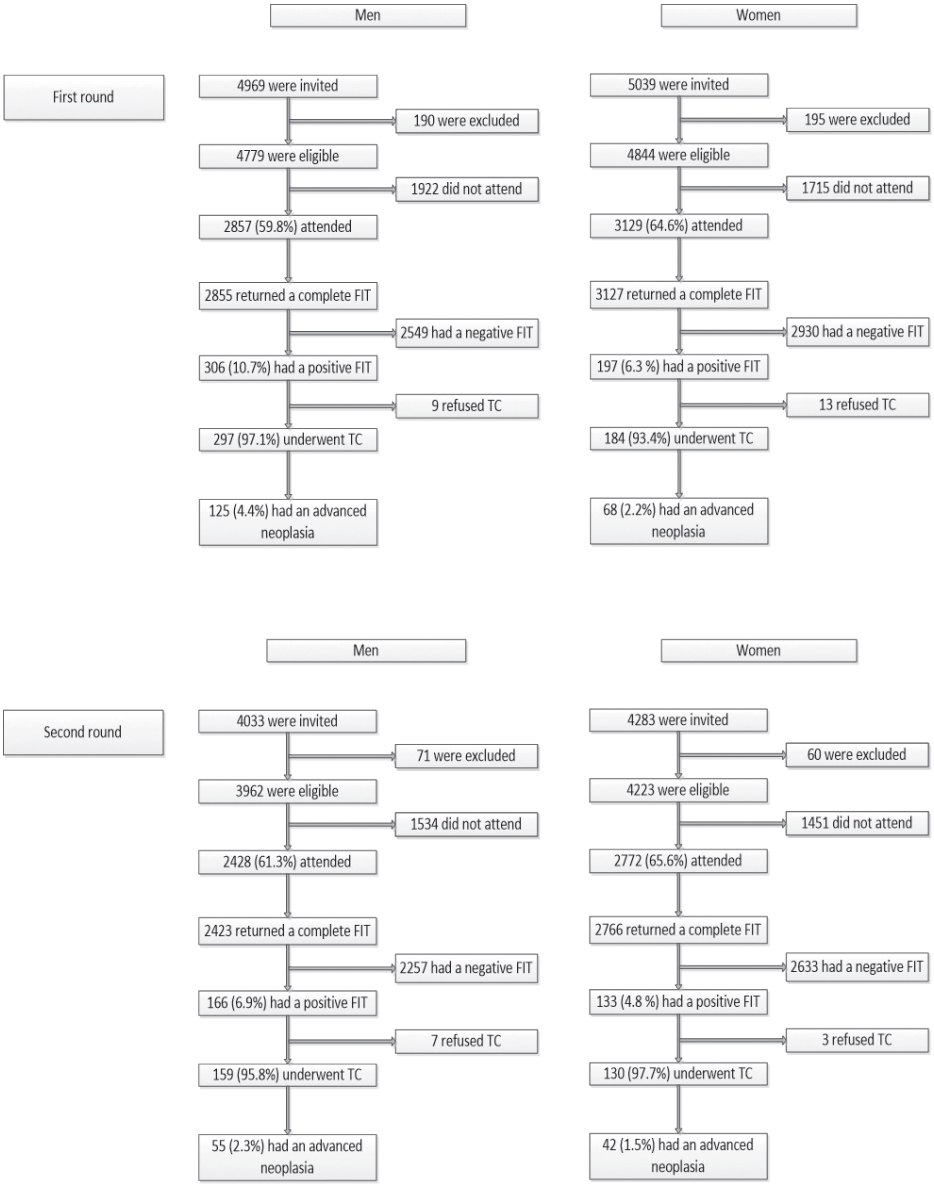
Statistical analysis

Differences in proportions between men and women for the test characteristics were analyzed by Chi-square testing. In case of more than two categorical variables, we changed to contingency table analyses [21]. Fecal hemoglobin concentrations were assessed in men and women. Differences between gender were analyzed using the Mann-Whitney U test, as the data were not normally distributed. The normality of the distribution of continuous variables was assessed using a normal Q-Q plot. The positivity rate (PR), positive predictive value (PPV) and detection rate (DR) were calculated and described as percentages with 95% confidence intervals (95% CI). The PR was defined as the proportion of participants having a positive FIT. The PPV depends on sensitivity and specificity, but also on the baseline prevalence of a disease in the population. Here, the PPV for detection of advanced neoplasia was defined as the number of subjects with advanced neoplasia divided by all FIT-positive screenees who underwent colonoscopy. Advanced neoplasia included CRC and advanced adenomas. An advanced adenoma was defined as an adenoma with a diameter ≥ 10 mm, and/or with a ≥ 25% villous component, and/or highgrade dysplasia. The DR was defined as the proportion of participants being diagnosed with advanced neoplasia divided by all screened individuals with an analyzable screening test. The number needed to scope (NNscope) describes the number of colonoscopies to find one screenee with an advanced neoplasia or CRC. The number needed to screen (NNscreen) was calculated as the number of complete screening tests needed to find one advanced neoplasia or CRC. All test characteristics were separately calculated for cut-off levels of 10, 15, 20, 25, 30, 35 and 40 µg Hb/g, respectively. FIT test characteristics and FIT concentrations were adjusted for age via logistic regression. True-positives were participants with a positive FIT result and advanced neoplasia detected during colonoscopy. False-positives were participants with a positive FIT result and non-advanced adenoma or no findings detected during colonoscopy. Likewise, the false-positive rate (FPR) was defined as subjects who had a positive FIT, but no advanced neoplasia on follow-up colonoscopy (i.e. only non-advanced adenoma, hyperplastic polyps or no findings at all), divided by the total number of screenees. All tests were conducted using SPSS version 20.0 and a *p*-value below 0.05 was considered statistically significant using 2-sided tests.

Ethical approval

The study was approved by the Dutch National Health Council and the Institutional Review Board of the Erasmus MC University Medical Centre (MEC-2005-264 and MEC-2008-029). All screenees gave written informed consent.

Figure 1: **Trial profile**



FIT: fecal immunochemical test; TC: total colonoscopy

RESULTS

The trial profile as described previously is summarized in Figure 1 [4, 20]. In total, 59.8% (95% CI: 58.4-61.2) of men and 64.6% (95% CI: 63.2-65.9) of women participated in the first round (p<0.001), and 61.3% (95% CI: 59.8-62.8) of men and 65.6% (95% CI: 64.2-67.1) of women participated in the second round (p<0.001), respectively.

Proportion of positive tests

In the first round, 306 male screenees (10.7%; 95% CI: 9.6-11.9%) and 197 female screenees (6.3%; 95% CI: 5.5-7.2%) tested positive at a cut-off level of 10 µg Hb/g (p<0.001). Men showed significantly higher positivity rates than women at the full range of FIT cut-off levels in the first round (Table 1). In the second round, 6.8% (95% CI: 5.9-7.9%) of men (n=166) and 4.8% (95% CI: 4.1-5.7%) of women (n=133) tested positive at a cut-off level 10 µg Hb/g (p=0.002). The proportion of positive tests remained significantly higher in males up to the cut-off level of 25 µg Hb/g. Above this cut-off level no significant differences were seen in positivity rates between both sexes in the second round (Table 1; Appendix 2). In both rounds gender was significantly associated with the positivity rate after adjusting for age.

Figure 2 shows the difference between men and women per FIT cut-off category in the first and second round. In the first round, 1422 men (49.8%) and 1656 women (53.0%) had a FIT result of 0 µg Hb/g. This was 1779 (72.8%) and 2096 (75.1%) in the second round, respectively. Men more often had hemoglobin levels of 10-20 µg Hb/g (3.9% vs. 2.4%, p=0.001), 20-30 µg Hb/g (1.3% vs. 0.6%, p=0.006), and ≥40 µg Hb/g (4.9% vs. 2.8%, p<0.001) in the first round compared to women. In the second round, men more often had hemoglobin levels of 20-30 µg Hb/g (1.2% vs. 0.6%, p=0.027).

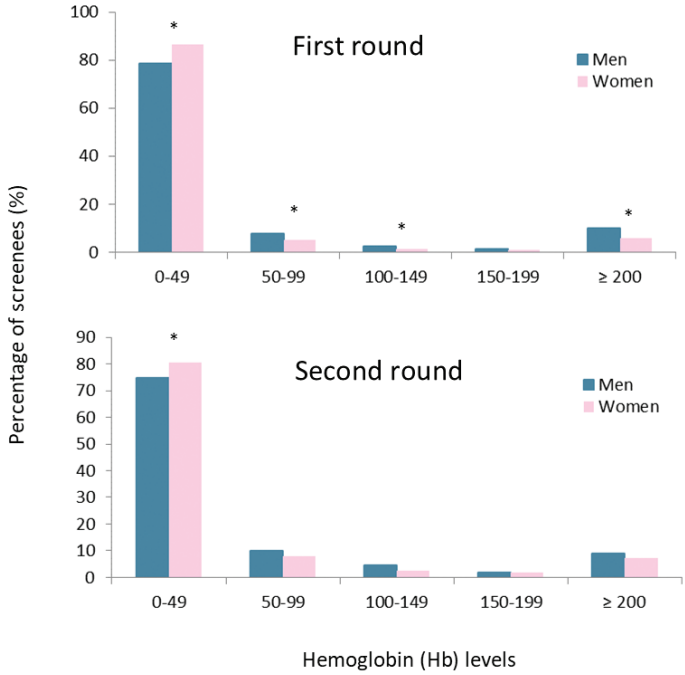
Table 1: **Test characteristics of FIT at different cut-off levels**

FIRST SCREENING ROUND															
	Positivity rate				Positive predictive value				Detection rate						
	Men		Women		Men		Women		Men				Women		
Cut-off					AN	CRC	AN	CRC	AN		CRC		AN		CRC
µg/ml	n	% (95% CI)	n	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)
50	306	10.7 (9.6-11.9)*	197	6.3 (5.5-7.2)	42 (37-48)	6.1 (4-9)	37 (30-44)	6.0 (3-11)	125	4.4 (3.7-5.2)*	18	0.6 (0.4-1.0)	68	2.2 (1.7-2.7)	11	0.4 (0.2-0.6)
75	228	8.0 (7.0-9.0)*	147	4.7 (4.0-5.5)	51 (45-58)*	7.7 (5-12)	40 (32-48)	7.2 (4-13)	113	4.0 (3.3-4.7)*	17	0.6 (0.4-1.0)	55	1.8 (1.4-2.3)	10	0.3 (0.2-0.6)
100	195	6.8 (6.0-7.8)*	122	3.9 (3.3-4.6)	55 (48-62)	9.0 (6-14)	44 (35-53)	8.7 (5-15)	104	3.6 (3.0-4.4)*	17	0.6 (0.4-1.0)	50	1.6 (1.2-2.1)	10	0.3 (0.2-0.6)
125	173	6.1 (5.2-7.0)*	112	3.6 (3.0-4.3)	57 (49-64)	8.9 (6-14)	46 (37-56)	9.4 (5-17)	95	3.3 (2.7-4.1)*	15	0.5 (0.3-0.9)	49	1.6 (1.2-2.1)	10	0.3 (0.2-0.6)
150	158	5.5 (4.8-6.4)*	103	3.3 (2.7-4.0)	60 (52-67)	9.8 (6-16)	48 (38-58)	9.2 (5-17)	91	3.2 (2.6-3.9)*	15	0.5 (0.3-0.9)	47	1.5 (1.1-2.0)	9	0.3 (0.1-0.6)
175	147	5.2 (4.4-6.0)*	92	2.9 (2.4-3.6)	60 (52-68)	9.9 (6-16)	49 (39-60)	9.2 (5-17)	85	3.0 (2.4-3.7)*	14	0.5 (0.3-0.8)	43	1.4 (1.0-1.8)	8	0.3 (0.1-0.5)
200	141	4.9 (4.2-5.8)*	88	2.8 (2.3-3.5)	60 (52-68)	9.6 (6-16)	51 (40-61)	9.6 (5-18)	82	2.9 (2.3-3.6)*	13	0.5 (0.3-0.8)	42	1.3 (1.0-1.8)	8	0.3 (0.1-0.5)
SECOND SCREENING ROUND															
	Positivity rate				Positive predictive value				Detection rate						
	Men		Women		Men		Women		Men				Women		
Cut-off					AN	CRC	AN	CRC	AN		CRC		AN		CRC
µg/ml	*n*	% (95% CI)	*n*	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)
50	166	6.8 (5.9-7.9)*	133	4.8 (4.1-5.7)	35 (28-42)	3.8 (2-8)	32 (25-41)	3.8 (2-9)	55	2.3 (1.7-2.9)*	6	0.2 (0.1-0.6)	42	1.5 (1.1-2.0)	5	0.2 (0.1-0.4)
75	131	5.4 (4.6-6.4)*	96	3.5 (2.8-4.2)	36 (28-45)	3.2 (1-8)	36 (27-46)	4.2 (2-11)	45	1.9 (1.4-2.5)	4	0.2 (0.1-0.4)	34	1.2 (0.9-1.7)	4	0.1 (0.1-0.4)
100	100	4.1 (3.4-5.0)*	79	2.9 (2.3-3.5)	44 (34-54)	4.3 (2-11)	37 (27-48)	5.1 (2-13)	41	1.7 (1.2-2.3)*	4	0.2 (0.1-0.4)	29	1.0 (0.7-1.5)	4	0.1 (0.1-0.4)
125	83	3.4 (2.8-4.2)*	68	2.5 (1.9-3.1)	47 (37-59)	5.1 (2-13)	36 (25-48)	4.5 (2-13)	37	1.5 (1.1-2.1)*	4	0.2 (0.1-0.4)	24	0.9 (0.6-1.3)	3	0.1 (0.0-0.3)
150	70	2.9 (2.3-3.6)	61	2.2 (1.7-2.8)	49 (37-60)	6.1 (2-15)	37 (26-50)	5.0 (2-14)	32	1.3 (0.9-1.9)	4	0.2 (0.1-0.4)	22	0.8 (0.5-1.2)	3	0.1 (0.0-0.3)
175	63	2.6 (2.0-3.3)	55	2.0 (1.5-2.6)	51 (38-63)	6.8 (2-17)	39 (27-52)	5.6 (2-16)	30	1.2 (0.9-1.8)	4	0.2 (0.1-0.4)	21	0.8 (0.5-1.2)	3	0.1 (0.0-0.3)
200	58	2.4 (1.9-3.1)	49	1.8 (1.3-2.3)	50 (37-63)	5.6 (2-16)	40 (27-54)	6.2 (2-18)	27	1.1 (0.8-1.6)	3	0.1 (0.0-0.4)	19	0.7 (0.4-1.1)	3	0.1 (0.0-0.3)

FIT: fecal immunochemical test; CRC: colorectal cancer; advanced neoplasia: adenoma ≥ 10mm, villous component (≥ 25% villous) or high-grade dysplasia; PPV: positive predictive value; NNscope: number needed to scope to detect one screenee with an advanced neoplasia; NNscreen: number needed to screen to detect one screenee with an advanced neoplasia

*P<0.05 Scores for men compared to scores for women in that particular FIT cut-off level

Figure 2: **Distribution of hemoglobin concentrations (µg Hb/ml) among FIT attendees per gender**



FIT: fecal immunochemical test; *: Significant difference in the percentage of screenees for that FIT cut-off category

Test characteristics

For both screening rounds, the uptake of colonoscopy among subjects with a positive FIT was high (round I: 97% of men and 93% of women, p=0.050; round II: 96% for both men and women, p=0.955, Figure 1).

In the first round, differences in PPV for advanced neoplasia between men and women were only significant at a cut-off level of 15 µg Hb/g (men: 51% (95% CI 45-58); women: 40% (95% CI: 32-48), p=0.032) (Table 1; Appendix 2). At higher cut-off levels, the PPV for advanced neoplasia tended to be higher in men, but these differences did not reach statistical significance. In the second round, no differences in PPV between men and women were observed. Likewise, the NNScope for advanced neoplasia and CRC were similar in both sexes. In the first round, the NNScope to find an advanced neoplastic lesion in men decreased from 2.4 (95% CI: 2.1-2.7) using a cut-off level of 10 µg Hb/g to 1.7 (95% CI: 1.5-1.9) at a cut-off level of 40 µg Hb/g. In women, the NNScope to find an advanced neoplastic lesion decreased from 2.7 (95% CI: 2.3-3.3) to 2.0 (95% CI: 1.6-2.5). In the second round, a similar pattern of decreasing NNscope was seen with increasing cut-offs. In both rounds gender was not significantly associated with

the PPV for advanced neoplasia after adjusting for age. A significantly higher FPR in men was found in both rounds at a cut-off level of 10 µg Hb/g: FPR round I: 6.3% in men vs. 4.1% in women, p<0.001; FPR round II: 4.6% in men vs. 3.3% in women, p=0.017). This difference remained significant until a cut-off level of 20 µg Hb/g in the first round, and a cut-off level of 15 µg Hb/g in the second round. Men showed higher detection rates of advanced neoplasia than women for the full range of FIT cut-off levels in the first round, and therefore the NNscreen to find an advanced neoplasia was significantly lower in men (Table 1). At a cut-off level of 10 µg Hb/g, the NNscreen to detect one subject with advanced neoplasia was 23 (95% CI: 19-27) in men and 46 (95% CI: 37-59) in women (p<0.001). In the second round, men also tended to have higher detection rates of advanced neoplasia compared to women, but these differences were only significant at cut-off levels of 10 10 µg Hb/g, 20 µg Hb/g, and 25 µg Hb/g, respectively (Table 1; Appendix 2). Likewise, a lower NNscreen to detect one advanced neoplasia was seen in men at these cut-off levels (cut-off 10 µg Hb/g: men 44 (95% CI: 34-59), women 66 (95% CI: 50-91), p=0.046; cut-off 20 µg Hb/g: men 59 (95% CI: 44-83), women 95 (95% CI: 67-143), p=0.045; cut-off 25 µg Hb/g: men 65 (95% CI: 48-91), women 115 (95% CI: 77-167), p=0.028, respectively). Gender was significantly associated with the DR of advanced neoplasia after adjusting for age in the first round, but not in the second round.

Fecal hemoglobin concentrations and true- and false-positivity

No differences were seen when comparing the fecal hemoglobin concentrations between true-positive men and women (65 µg Hb/g (IQR 24; 196) vs 72 µg Hb/g (IQR 29; 211), p=0.840) and false-positive men and women (23 µg Hb/g (IQR 14; 65) vs 24 µg Hb/g (IQR 14; 58), p=0.647) for the first and second round combined.

DISCUSSION

Information on gender differences in population-based FIT screening was limited until now. This study, in which conclusions were based on a large number of screening-naïve men and women in a two-round FIT screening setting, provides insight in this matter. We observed higher positivity rates in men at the full range of cut-off levels. This was reflected by higher true-positive rates (detection rates) and higher false-positive rates (FPR). Likewise, the number needed to screen was lower in men for all cut-off levels. A higher PPV for advanced neoplasia in men was only seen at a cut-off level of 15 µg Hb/g in the first round. Data on the performance of FIT in men and women are of key importance given the current widespread use of FIT as primary screening tool.

Similar differences in detection rates of advanced neoplasia between both sexes were found in two colonoscopy screening studies [11, 12]. The higher detection rate is

related to a higher prevalence of advanced lesions in men. Since negative screenees did not undergo a colonoscopy in our study, we were unable to calculate the FIT sensitivity and specificity. However, the relative difference in detection rates of advanced neoplasia between men and women in our study is higher than what one would expect based on the relative risk for developing CRC in the screening age group. Per 100.000 inhabitants, 1.773/100.000 men and 1.050/100.000 women aged 50-74 years were diagnosed with CRC in the Netherlands in 2013 (Comprehensive Cancer Centre the Netherlands). The incidence rate is therefore 1.7-fold higher in men. We found a two-fold higher detection rate of advanced neoplasia in men in the first round. Although it should be interpreted with care since the data is rather speculative and only offers a suggestion, this may indicate a higher FIT sensitivity in men. This is in line with two colonoscopy screening studies where subjects received a FOBT prior to colonoscopy (gFOBT or FIT). Both reported a higher test sensitivity in men [18, 19]. Furthermore, the higher FPR in men may be the result of a lower test specificity. Specificity is defined by the proportion of people without the disease that also test negative. We do not know the exact number of people without disease (advanced neoplasia) since people with a negative FIT did not undergo colonoscopy. However, given the higher underlying prevalence of advanced neoplasia in men, the number of men without advanced neoplasia will consequently be lower than the number of women. Therefore, the higher number of male screenees with a false-positive test indicates that the FIT specificity is lower in men. This is in line with the results of the two aforementioned colonoscopy screening studies [18, 19]. In addition, we calculated the FPR for the scenario in which subjects with a positive FIT, who had no adenoma or CRC at follow-up colonoscopy (i.e. only hyperplastic polyps or no findings at all). After including also the non-advanced adenomas, we did not see any differences between men and women for the different cut-off levels (Appendix 3). This would imply that the higher FPR in men is mainly caused by positive FITs due to detection of non-advanced adenomas. Our finding of similar positive predictive values of FIT in men and women contrasts with a German study on the performance of one guaiac and several immunochemical fecal occult blood tests. In this study men had substantially higher positive predictive values than women at any FIT cut-off point [22].

The key question in the interpretation of these findings is whether and to what extent the observed gender differences are of clinical and/or public health relevance. Some studies suggest the cut-off should differ between men and women to reach the same FIT sensitivity in men and women. However, we think it is better to determine the optimal cut-off by other measures, in particular PPV, since the PPV is a measure for efficient use of colonoscopy resources, and also for the individual reflects the chance that unnecessary harm is done. As screening colonoscopies are performed on healthy individuals, the number of unnecessary colonoscopies must be brought to an

absolute minimum. This is even more relevant since all colonoscopies carry a small risk of serious complications, such as bleeding and perforation [23, 24]. In addition, colonoscopy capacity in several countries is limited and costly. The higher FPR in men in both screening rounds indicates that a significantly larger number of men underwent follow-up colonoscopy and did not have advanced neoplasia. However, the chance that a colonoscopy is unnecessary after a positive FIT is equal in men and women, which is demonstrated by the similar PPV at a cut-off level of 10 µg Hb/g. Therefore, one could argue not to change cut-off values in men and women. In theory, if the same differences would persist between men and women in a larger sample, the difference in PPV would become significant. The PPV could then be improved by a higher cut-off in women, but this would be at the expense of the NNscreen in women. Optimal cut-off values for men and women can further be determined by taking other major determinants into account, including the incidence of neoplasia, the life expectancy, the intended screening interval, and cost-effectiveness. This can be realized using the current data combined with a microsimulation model [25-27]. The resulting information will be of great value, since FIT screening is expected to become current practice in more and more countries in the upcoming years. We were not able to determine possible differences between subgroups, e.g. investigating the PPV in men and women for different age groups. Information regarding subgroups would be of great value, as it might help to improve CRC risk-stratification based screening. A recent study explored the potential gains of using a risk prediction model in CRC screening (Stegeman et al, Gut 2014). Subjects aged 50-75 years who were invited to undergo colonoscopy, were asked to perform a FIT and to complete a risk questionnaire prior to colonoscopy. Based on the questionnaire data and the FIT results, a multivariable risk model was developed which included the following factors: total calcium intake, family history, age and FIT result. Combining risk stratification with the FIT result showed better accuracy than screening with only FIT, with better sensitivity at similar specificity levels, and more cases of advanced adenoma detected with a similar number of colonoscopies. Clearly, risk stratification can be used as a tool to improve the effectiveness of screening. Future studies should evaluate the practical implications of pre-selection with a risk algorithm, with a focus on costs and participation rate.

Some limitations must be acknowledged. As already mentioned above, it was not possible to explicitly estimate sensitivity and specificity, because negative screenees did not undergo colonoscopy. Second, different screening intervals were applied in the second round. However, these intervals did not influence the results, since detection rates and positive predictive values of advanced neoplasia were comparable for the different intervals [20]. Furthermore, perhaps if our study population had been larger, differences in PPV would have become significant for all cut-off levels, indicating a better

test performance in men. Finally, we tried to determine gender differences between proximal and distal advanced lesions, but our numbers were too small to consider for this manuscript.

In conclusion, this population-based trial provides important data on performance of one-sample FIT screening in men and women at different cut-off levels. Men have higher positivity rates than women, reflected by both higher detection rates and a higher FPR. A higher FPR in men implies that specificity is lower in men than in women. Positive predictive values did not differ significantly for most cut-off levels. The resulting harm-to-benefit ratio, reflected in the positive predictive value, did not differ. Therefore, the use of similar cut-off values in men and women in a FIT screening setting seems reasonable.

REFERENCES

1. Hewitson, P., P. Glasziou, L. Irwig, et al., *Screening for colorectal cancer using the faecal occult blood test, Hemoccult.* Cochrane Database of Systematic Reviews, 2007(1).
2. Allison, J.E., L.C. Sakoda, T.R. Levin, et al., *Screening for colorectal neoplasms with new fecal occult blood tests: update on performance characteristics.* J Natl Cancer Inst, 2007. **99**(19): p. 1462-70.
3. Park, D.I., S. Ryu, Y.H. Kim, et al., *Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening.* Am J Gastroenterol, 2010. **105**(9): p. 2017-25.
4. Hol, L., M.E. van Leerdam, M. van Ballegooijen, et al., *Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy.* Gut, 2010. **59**(1): p. 62-68.
5. Imperiale, T.F., *Noninvasive screening tests for colorectal cancer.* Dig Dis, 2012. **30 Suppl 2**: p. 16-26.
6. Rabeneck, L., R.B. Rumble, F. Thompson, et al., *Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening.* Can J Gastroenterol, 2012. **26**(3): p. 131-47.
7. Halloran, S.P., G. Launoy, M. Zappa, et al., *European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing.* Endoscopy, 2012. **44 Suppl 3**: p. SE65-87.
8. European Colorectal Cancer Screening Guidelines Working, G., L. von Karsa, J. Patnick, et al., *European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication.* Endoscopy, 2013. **45**(1): p. 51-9.
9. Kuipers, E.J., T. Rosch, and M. Bretthauer, *Colorectal cancer screening--optimizing current strategies and new directions.* Nat Rev Clin Oncol, 2013. **10**(3): p. 130-42.
10. Brenner, H., M. Hoffmeister, C. Stegmaier, et al., *Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840 149 screening colonoscopies.* Gut, 2007. **56**(11): p. 1585-1589.
11. Regula, J., M. Rupinski, E. Kraszewska, et al., *Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia.* New England Journal of Medicine, 2006. **355**(18): p. 1863-1872.
12. Schoenfeld, P., B. Cash, A. Flood, et al., *Colonoscopic screening of average-risk women for colorectal neoplasia.* New England Journal of Medicine, 2005. **352**(20): p. 2061-2068.
13. Lieberman, D.A., D.G. Weiss, W.V. Harford, et al., *One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon.* New England Journal of Medicine, 2001. **345**(8): p. 555-560.
14. Garcia, M., N. Mila, G. Binefa, et al., *False-positive results from colorectal cancer screening in Catalonia (Spain), 2000-2010.* J Med Screen, 2012. **19**(2): p. 77-82.
15. Steele, R.J., I. Kostourou, P. McClements, et al., *Effect of gender, age and deprivation on key performance indicators in a FOBT-based colorectal screening programme.* J Med Screen, 2010. **17**(2): p. 68-74.
16. Weller, D., D. Coleman, R. Robertson, et al., *The UK colorectal cancer screening pilot: results of the second round of screening in England.* Br J Cancer, 2007. **97**(12): p. 1601-5.
17. Steele, R.J., P. McClements, C. Watling, et al., *Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site.* Gut, 2012. **61**(4): p. 576-81.

18. Brenner, H., U. Haug, and S. Hundt, *Sex Differences in Performance of Fecal Occult Blood Testing.* American Journal of Gastroenterology, 2010. **105**(11): p. 2457-2464.
19. Stegeman, I., T.R. de Wijkerslooth, E.M. Stoop, et al., *Risk factors for false positive and for false negative test results in screening with fecal occult blood testing.* Int J Cancer, 2013. **133**(10): p. 2408-14.
20. van Roon, A.H., S.L. Goede, M. van Ballegooijen, et al., *Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening.* Gut, 2013. **62**(3): p. 409-15.
21. Bosman, F.T., Carneiro F., Hruban R.H., et al., *WHO Classification of Tumours of the Digestive System, Fourth Edition.* 2010.
22. Brenner, H., U. Haug, and S. Hundt, *Sex differences in performance of fecal occult blood testing.* Am J Gastroenterol, 2010. **105**(11): p. 2457-64.
23. Panteris, V., J. Haringsma, and E.J. Kuipers, *Colonoscopy perforation rate, mechanisms and outcome: from diagnostic to therapeutic colonoscopy.* Endoscopy, 2009. **41**(11): p. 941-51.
24. Nelson, D.B., K.R. McQuaid, J.H. Bond, et al., *Procedural success and complications of large-scale screening colonoscopy.* Gastrointest Endosc, 2002. **55**(3): p. 307-14.
25. Lansdorp-Vogelaar, I., K.M. Kuntz, A.B. Knudsen, et al., *Contribution of screening and survival differences to racial disparities in colorectal cancer rates.* Cancer Epidemiol Biomarkers Prev, 2012. **21**(5): p. 728-36.
26. Wilschut, J.A., J.D. Habbema, M.E. van Leerdam, et al., *Fecal occult blood testing when colonoscopy capacity is limited.* J Natl Cancer Inst, 2011. **103**(23): p. 1741-51.
27. Wilschut, J.A., L. Hol, E. Dekker, et al., *Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening.* Gastroenterology, 2011. **141**(5): p. 1648-55 e1.

SUPPLEMENTARY MATERIAL

Appendix 1. **The FITTER Check-List for the Reporting of Studies Using FITs for Hemoglobin**

Specimen Collection and Handling

The name of the specimen collection device and supplier (address):

- Supplier is Eiken Chemical Co.
- The name of the collection device is S-bottle (OC- Sensor Micro Sampling Bottle, Eiken Chemical Co).

Description of specimen collection device (vial with probe/stick, card, other):

- S-bottle with a stick in the shell and buffer in the tube. After collection, the stick has to be put back in the test tube. The tip of the stick (with feces) ends in the buffer. The buffer stabilizes the hemoglobin in the feces sample.

Description of specimens used if an in vivo study (single or pooled feces, artificial matrix with added blood, and so forth):

- Single feces.

Details of fecal collection method (sampling tech- nique and number of samples):

- Single feces sample of 1 bowel movement.

Sampling technique:

- Stab the stick 4 times in the feces, afterward put the stick back into the test tube.

Who collected the specimens from the samples (patient, technician, and so forth)?

- Participant (patient) at home.

Number of fecal specimens used in the study (single, pooled, individual patient feces):

- One single sample per participant.

Mean mass of feces collected.* Essential volume of buffer into which specimen is taken by probe, applicator stick, or card*:

- 10 mg feces was collected in 2 mL buffer.

Time and storage conditions of fecal specimen from “passing” to sampling, including time and temperature (median and range):

- In our opinion, the feces sampling is direct after passing the stool (within 15 minutes), returned by mail the same day, we advise the patient to store the tube (properly packed) in the refrigerator.

Time and storage of collection devices from specimen collection to analysis, including time and temperature (median and range). A concise description of the process from collection to analysis is recommended:

- After collection of the feces sample, the participant sends the collection device to the laboratory in a return box by mail.

- The laboratory checks the collection date of the test (at the S bottle or at informed consent) and the date of entrance in the laboratory.
- If the return time is 7 days or fewer between the collection date and the entrance date: all right
- If the return time is longer than 7 days between the collection date and the entrance date: a comment is made in the database.
- All samples that entered the laboratory were analyzed in the laboratory. In addition, samples with a return time longer than 7 days.
    a) If return time was longer than 7 days and the test was positive (Hb level, ≤10 µg Hb/g feces), then the participant is referred to colonoscopy.
    b) The reason for this is that the positivity of the test has a higher importance than the return time of the sample.
    c) If return time is longer than 7 days and the test is negative (Hb, ≤10 mg Hb/g feces), the participant receives a new test. The reason for this is that during this long return time the Hb could have degraded. The participant did not receive a result mail from the first (overtime) test. The participant receives a result mail from the second test, when the return time is fewer than 7 days.

After the test entered the laboratory, trained laboratory personnel entered the test result into the database correctly. The test was stored at -20°C until analysis, for at most 14 days. At least 95% of the tests were analyzed within 7 days after entering the laboratory.

Analysis

Name of analyzer, model, supplier (address), number of systems if more than one was used:

- OC Sensor m (micro) system. Supplier: Eiken Chemical Co.
- Address: 4-19-9 Taito, Taitoku, Tokyo, 110-8408, Japan.

Number of times each sample was analyzed. Essential analytical working range* and whether samples outside this range were diluted (factor) and re-assayed:

- Samples were analyzed once.

Samples outside the range were not diluted or re- assayed. Outside the range was considered far above the highest calibration (488 µg Hb/g feces) and our cut-off level (≤10 µg Hb/g feces).

Source of calibrator(s) (supplier with address), number of calibrator(s), how concentrations were assigned,*and details of calibration process including frequency.

Calibrator, control low and control high are from Eiken Chemical Co.:

- Calibrator 1: 0 µg Hb/g

- Calibrator 2: 12 µg Hb/g (10% difference accepted)
- Calibrator 3: 24 µg Hb/g (5% difference accepted)
- Calibrator 4: 49 µg Hb/g (5% difference accepted)
- Calibrator 5: 98 µg Hb/g (5% difference accepted)
- Calibrator 6: 195 µg Hb/g (5% difference accepted)
- Control low: between 24 and 36 µg Hb/g, dependent on lot number
- Control high between 111 and 167 µg Hb/g, dependent on lot number

The calibration process occurred every week. After running calibration and controls low and high, and all values were within the margins, the test samples were run for analysis.

Analytical imprecision,* ideally with the number of samples analyzed, concentrations, and mean, SD, and co-efficient of variation:

- We measured the low and high control from one lot number because the controls have a standard con- centration. For the low control the variation coefficient is between 5% and 11%, for the high control the variation coefficient is between 7.5% and 12.7%.
- The limit of quantitation was detected at 12 µg Hb/g feces, with a 95% CI and co-efficient of variation of 20%.

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 analytic runs.

Analytic runs were accepted only if the calibration and controls were in the margins.

Participation in external quality assessment schemes: (name and address of scheme), frequency of challenges, performance attained:

- External assessment: once a year 2 external samples from Eiken, these had to be measured 3 times and results were sent to Eiken for evaluation. Since 2014 it is possible to take part in an external quality assessment scheme of Foundation for Quality Medical Laboratory Diagnostics in The Netherlands, 6 samples have to be measured every 3 months.
- Accreditation held by the analytic facility (address). Desirable for laboratory evaluations.
- No accreditation available.
- The number, training, and expertise of the persons performing the analyses and recording the results.
- The analyses were performed by 4 laboratory analysts, 2 people recorded the results.

- The Dutch supplier trained the analysts at the start of the first screening round.

Result Handling

Mode of data collection: manual recording or via automatic download to information technology system, single or double reading:

- Automatic download to IT system, single reading.

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

- 50 ng/mL Hb / 10 µg Hb/g feces.

Cut-off concentration(s) if used and explanation of how assigned locally or assigned by the manufacturer*:

- 10 µg Hb/g feces, assigned locally. When we started the first round of screening in 2006, different cut-off levels were used in Italy and Japan. We have chosen for a low cut-off level because we did not have an idea what would be the best cut-off level in The Netherlands in the asymptomatic population. During the different rounds of screening we did not change the cut-off level.

Were the analysts blinded (masked) to the results of the reference investigation and other clinical information?

- After analysis of the tests, the analysts were not blinded to the name, address, or date of birth of the participant. However, analysts were blinded for other clinical information.

Appendix 2. **Positivity Rates and Detection Rates of FIT at Different Cut-Off Levels**

Cut-off µg Hb/g	Positivity rate					Detection rate AN					Detection rate CRC				
	Men		Women			Men		Women			Men		Women		
	n	% (95% CI)	n	% (95% CI)	*P* value	n	% (95% CI)	n	% (95% CI)	*P* value	n	% (95% CI)	n	% (95% CI)	*P* value
First screening round															
10	306	10.7 (9.6–11.9)[a]	197	6.3 (5.5–7.2)	<.001	125	4.4 (3.7–5.2)[a]	68	2.2 (1.7–2.7)	<.001	18	0.6 (0.4–1.0)	11	0.4 (0.2–0.6)	.121
15	228	8.0 (7.0–9.0)[a]	147	4.7 (4.0–5.5)	<.001	113	4.0 (3.3–4.7)[a]	55	1.8 (1.4–2.3)	<.001	17	0.6 (0.4–1.0)	10	0.3 (0.2–0.6)	.113
20	195	6.8 (6.0–7.8)[a]	122	3.9 (3.3–4.6)	<.001	104	3.6 (3.0–4.4)[a]	50	1.6 (1.2–2.1)	<.001	17	0.6 (0.4–1.0)	10	0.3 (0.2–0.6)	.113
25	173	6.1 (5.2–7.0)[a]	112	3.6 (3.0–4.3)	<.001	95	3.3 (2.7–4.1)[a]	49	1.6 (1.2–2.1)	<.001	15	0.5 (0.3–0.9)	10	0.3 (0.2–0.6)	.219
30	158	5.5 (4.8–6.4)[a]	103	3.3 (2.7–4.0)	<.001	91	3.2 (2.6–3.9)[a]	47	1.5 (1.1–2.0)	<.001	15	0.5 (0.3–0.9)	9	0.3 (0.1–0.5)	.147
35	147	5.2 (4.4–6.0)[a]	92	2.9 (2.4–3.6)	<.001	85	3.0 (2.4–3.7)[a]	43	1.4 (1.0–1.8)	<.001	14	0.5 (0.3–0.8)	8	0.3 (0.1–0.5)	.135
40	141	4.9 (4.2–5.8)[a]	88	2.8 (2.3–3.5)	<.001	82	2.9 (2.3–3.6)[a]	42	1.3 (1.0–1.8)	<.001	13	0.5 (0.3–0.8)	8	0.1 (0.1–0.5)	.193
Second screening round															
10	166	6.8 (5.9–7.9)[a]	133	4.8 (4.1–5.7)	.002	55	2.3 (1.7–2.9)[a]	42	1.5 (1.1–2.0)	.046	6	0.2 (0.1–0.6)	5	0.2 (0.1–0.4)	.601
15	131	5.4 (4.6–6.4)[a]	96	3.5 (2.8–4.2)	.001	45	1.9 (1.4–2.5)	34	1.2 (0.9–1.7)	.065	4	0.2 (0.1–0.4)	4	0.1 (0.1–0.4)	.851
20	100	4.1 (3.4–5.0)[a]	79	2.9 (2.3–3.5)	.012	41	1.7 (1.2–2.3)[a]	29	1.0 (0.7–1.5)	.045	4	0.2 (0.1–0.4)	4	0.1 (0.1–0.4)	.851
25	83	3.4 (2.8–4.2)[a]	68	2.5 (1.9–3.1)	.039	37	1.5 (1.1–2.1)[a]	24	0.9 (0.6–1.3)	.028	4	0.2 (0.1–0.4)	3	0.1 (0.0–0.3)	.579
30	70	2.9 (2.3–3.6)	61	2.2 (1.7–2.8)	.118	32	1.3 (0.9–1.9)	22	0.8 (0.5–1.2)	.063	4	0.2 (0.1–0.4)	3	0.1 (0.0–0.3)	.579
35	63	2.6 (2.0–3.3)	55	2.0 (1.5–2.6)	.141	30	1.2 (0.9–1.8)	21	0.8 (0.5–1.2)	.081	4	0.2 (0.1–0.4)	3	0.1 (0.0–0.3)	.579
40	58	2.4 (1.9–3.1)	49	1.8 (1.3–2.3)	.116	27	1.1 (0.8–1.6)	19	0.7 (0.4–1.1)	.101	3	0.1 (0.0–0.4)	3	0.1 (0.0–0.3)	.871

AN, advanced colorectal neoplasia: adenoma 10 mm, villous component (25% villous) or high-grade dysplasia.

[a] *P* < .05 scores for men compared with scores for women for that particular FIT cut-off level.

Appendix 3. **Positive Predictive Values of FIT at Different Cut-Off Levels**

	Round 1					Round 2				
Cut-off	Men		Women			Men		Women		
µg *Hb/g*	n	% (95% CI)	n	% (95% CI)	P-value	n	% (95% CI)	n	% (95% CI)	P-value
First screening round										
10	125	42 (37–48)	68	37 (30–44)	0.265	18	6.1 (4–9)	11	6.0 (3–11)	0.971
15	113	51 (45–58)a	55	40 (32–48)	0.032	17	7.7 (5–12)	10	7.2 (4–13)	0.861
20	104	55 (48–62)	50	44 (35–53)	0.051	17	9.0 (6–14)	10	8.7 (5–15)	0.929
25	95	57 (49–64)	49	46 (37–56)	0.096	15	8.9 (6–14)	10	9.4 (5–17)	0.887
30	91	60 (52–67)	47	48 (38–58)	0.074	15	9.8 (6–16)	9	9.2 (5–17)	0.871
35	85	60 (52–68)	43	49 (39–60)	0.123	14	9.9 (6–16)	8	9.2 (5–17)	0.869
40	82	60 (52–68)	42	51 (40–61)	0.16	13	9.6 (6–16)	8	9.6 (5–18)	0.984
Second screening round										
10	55	35 (28-42)	42	32 (25-41	0.683	6	3.8 (2-8)	5	2.3 (1.7–2.9)	0.974
15	45	36 (28–45)	34	36 (27–46)	0.974	4	3.2 (1–8)	4	1.9 (1.4–2.5)	0.692
20	41	44 (34–54)	29	37 (27–48)	0.392	4	4.3 (2–11)	4	1.7 (1.2–2.3)a	0.787
25	37	47 (37–59)	24	36 (25–48)	0.158	4	5.1 (2–13)	3	1.5 (1.1–2.1)a	0.855
30	32	49 (37–60)	22	37 (26–50)	0.181	4	6.1 (2–15)	3	1.3 (0.9–1.9)	0.795
35	30	51 (38–63)	21	39 (27–52)	0.202	4	6.8 (2–17)	3	1.2 (0.9–1.8)	0.787
40	27	50 (37–63)	19	40 (27–54)	0.291	3	5.6 (2–16)	3	1.1 (0.8–1.6)	0.882

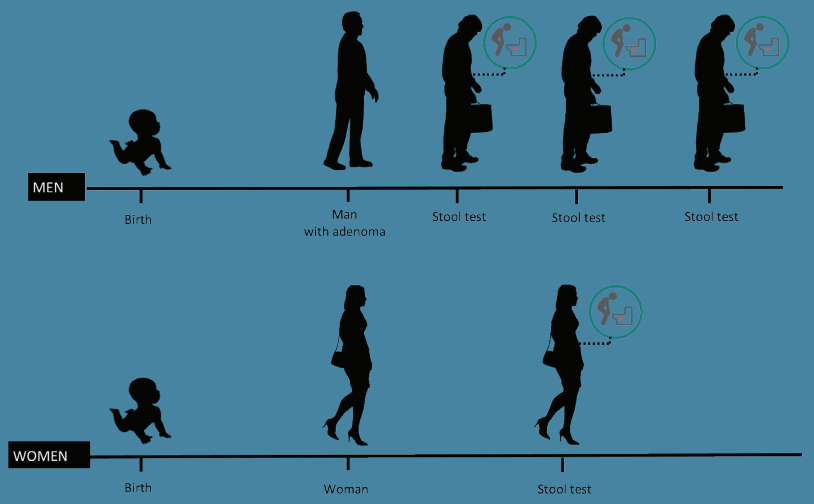
AN, advanced colorectal neoplasia: adenoma ≥10 mm, villous component (≥25% villous) or high-grade dysplasia.

[a]*P* < .05 scores for men compared with scores for women for that particular FIT cut-off level.

Appendix 4. **FPR for FIT participants over both screening rounds.**

	Round 1			Round 2		
Cut-off level	Men	Women		Men	Women	
µg *Hb/g*	%	%	P	%	%	P
10	2.9	3.4	0.194	2.3	2.5	0.633
15	2.1	2	0.89	1.6	2	0.244
20	1.5	1.6	0.898	1.2	1.4	0.588
25	1.3	1.3	0.947	1.1	1	0.836
30	1.2	1.1	0.722	0.9	0.8	0.741
35	1	1	0.967	0.8	0.7	0.697
40	0.9	1	0.942	0.7	0.7	0.786

NOTE. The FPR was for the scenario in which subjects with a positive FIT had no adenoma or CRC at follow-up colonoscopy (ie, only hyperplastic polyps or no findings at all.


MEN
Birth
Man
with adenoma
Stool test
Stool test
Stool test
WOMEN
Birth
Woman
Stool test

Chapter 3

Do men and women need to be screened differently with faecal immunochemical testing? A cost-effectiveness analysis.

Miriam P. van der Meulen[1], Atija Kapidzic[2] , Monique E. van Leerdam[4] , Alex van der Steen[1] , Ernst J. Kuipers[2,3], Manon C. Spaander[2], Harry J. de Koning[1], Lieke Hol[2] , Iris Lansdorp-Vogelaar[1]

Cancer Epidemiol Biomarkers Prev. 2017 Aug;26(8):1328-1336. doi: 10.1158/1055-9965. EPI-16-0786. Epub 2017 May 17.

ABSTRACT

Background:

Several studies suggest test characteristics for the faecal immunochemical test (FIT) differ by gender, triggering a debate whether men and women should be screened differently. We used the microsimulation model MISCAN-Colon to evaluate whether screening stratified by gender is cost-effective.

Methods:

We estimated gender-specific FIT characteristics based on first round positivity and detection rates observed in a FIT screening pilot (CORERO-1). Subsequently, we used the model to estimate harms, benefits and costs of 480 gender-specific FIT screening strategies and compared them with uniform screening.

Results:

Biennial FIT screening from age 50-75 was less effective in women than men (35.7 versus 49.0 QALYs gained, respectively) at higher costs (€42,161 versus -€5,471 respectively). However, the incremental QALYs gained and costs and of annual screening compared to biennial screening were more similar for both gender (8.7 QALYs gained and €26,394 for women versus 6.7 QALYs gained and €20,863 for men). Considering all evaluated screening strategies, optimal gender-based screening yielded at most 7% more QALYs gained than optimal uniform screening and even resulted in equal costs and QALYs gained from a willingness-to-pay threshold of €1300.

Conclusions:

FIT screening is less effective in women, but the incremental cost-effectiveness is similar in men and women. Consequently, screening stratified by gender is not more cost-effective than uniform FIT screening.

Impact:

Our conclusions support the current policy of uniform FIT screening.

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. Consequently, several countries and local initiatives across the world have adopted population-based screening for CRC. The majority of these initiatives have opted for some form of faecal immunochemical testing (FIT).(2, 3)

These screening programmes use the same approach for both genders despite age and gender disparities in prevalence of colorectal neoplasia and a higher life expectancy in women. We previously showed that a uniform approach is cost-effective for primary colonoscopy screening, because the lower prevalence of advanced neoplasia in women is compensated by a higher life expectancy.(4) However, we assumed that test characteristics for colonoscopy did not differ between genders, while there are strong indications that the test characteristics for FOBT (including FIT) differ between men and women.

A Scottish gFOBT screening study reported more screen-detected CRCs in men compared to women, whereas the number of interval CRCs was similar in both groups, suggesting higher sensitivity in men.(5) Two other studies likewise found a higher sensitivity of FIT for advanced neoplasia in men compared to women,(6, 7) and also a higher positive predictive value (PPV), whereas FIT specificity in men was found to be significantly lower.(6) The lower specificity in men was confirmed in the FIT screening trial CORERO-1 with a higher false positive rate in men compared to women.(8)

These studies triggered a debate whether men and women should be screened differently with FIT. For instance, lowering the FIT cut-off in women will increase their sensitivity towards that of men, or in contrast, increasing the cut-off in women will increase their PPV towards that of men.(9) Differences in test characteristics might affect the optimal cut-off for a positive FIT, but might also affect the optimal screening age range and interval. Microsimulation modelling can take these gender differences in test characteristics, but also in life expectancy and CRC incidence into account and estimate costs and quality adjusted life years (QALYs) gained of various screening strategies. In this study, the micro-simulation model MISCAN-Colon was used to determine optimal screening strategies for men and women and to study if screening stratified by gender is beneficial in terms of cost-effectiveness.

MATERIALS AND METHODS

We developed two separate versions of the microsimulation model MISCAN-Colon for men and women. We estimated sensitivity and specificity of FIT based on the

gender-specific positivity and detection rates observed in the CORERO-1 trial. We then simulated male and female populations screened with various FIT screening strategies to estimate the QALYs and costs of FIT screening by gender and determined efficient screening strategies for men and women. Thereafter we compared costs and effects of screening stratified by gender with uniform screening.

The CORERO trial

The CORERO-1 trial was a randomised controlled trial comparing attendance and detection rates of gFOBT, FIT and sigmoidoscopy at first round screening. For the current study we only used the data of FIT screening. Details from this trial have been described elsewhere.(10, 11) In short, screening-naïve subjects aged 50-74 years, living in the southwest of the Netherlands were selected through municipal population registers. Screenees assigned in the FIT study arm received a kit with a single FIT (OC-Sensor, Eiken, Japan). A cut-off of 10 µg haemoglobin/g faeces (equivalent to 50 ng haemoglobin/ml) was used to indicate a positive test result. This was followed by the recommendation for a diagnostic colonoscopy. In total, 4969 men and 5039 women were invited of which 59.8% of men and 64.6% of women returned the test. The positivity rate and detection rates were higher among men compared to women (Table 1). The positive predictive value (PPV) for advanced neoplasia did not differ significantly for men (42.1%) and women (37.0%) (p=0.265). Positivity rates, detection rates and the PPV at higher cut-offs can be found elsewhere.(8)

MISCAN-Colon

MISCAN-Colon is a well-established microsimulation model for CRC developed at Erasmus MC, the Netherlands. The model has been extensively described previously (12, 13) and in the Model Appendix in this thesis. In brief, MISCAN-Colon simulates life histories of a large group of individuals from birth to death. As each simulated person ages, one or more adenomas may develop. These adenomas can progress in size from small (≤5mm) to medium (6-9mm) to large size (≥10mm). Some adenomas can develop into preclinical cancer, which may progress through stages I to IV.

At any time during the development of the disease, the process may be interrupted because a person dies of other causes. With screening, CRC may be prevented by the detection and removal of adenomas or detected at an earlier stage with a more favourable survival. In this way, CRC incidence and/or CRC-related mortality can be reduced. The life years gained by screening are calculated as the difference in model-predicted life years lived in the population with and without CRC screening.

Model input

Natural history

We developed two versions of the MISCAN-Colon model, one for each gender. The two versions were separately calibrated to gender-specific pre-screening data on the age-specific incidence of CRC as observed in the Netherlands before the introduction of screening (between 1999 and 2003)(14) and the gender and age-specific prevalence and multiplicity distribution of adenomas as observed in autopsy studies.(15-22) The size distribution of adenomas was calibrated to the size distribution of adenomas detected in a colonoscopy trial.(23) Survival after clinical diagnosis is based on 1989-2003 survival data obtained from the Dutch Comprehensive Cancer Center.(14) The preclinical duration of CRC and the adenoma dwell-time were calibrated to the rates of interval and surveillance-detected cancers observed in randomised controlled trials evaluating screening using guaiac-based faecal occult blood tests and a once-only sigmoidoscopy.(24) The model outcomes showed good concordance with trial results (Supplementary Model Appendix).

Study population

We modelled the age distribution and life-expectancy separately for the male and female version of the model. We modelled the Dutch population aged 25 to 85 years in 2015(25) and all individuals were followed until death. Life-expectancy was based on gender-specific life-tables from 2011 obtained from Statistics Netherlands.(25)

Screening strategies

FIT screening was simulated in the population starting in year 2015. Individuals were offered screening according to different FIT screening schedules varying by:

- Age to start screening: 40, 45, 50, 55, 60 and 65 years
- Age to stop screening: 70, 75, 80 and 85 years
- Screening interval: 1, 1.5, 2 and 3 years

The cut-off level for a positive FIT result varied between 10 (FIT^{10}), 15, 20, 30 and 40 µg haemoglobin/g faeces (equivalent to 50-200 ng Hb/ml). This resulted in a total of 480 different screening strategies per gender.

If adenomas were detected, individuals entered a surveillance programme according to the Dutch guidelines for follow-up after polypectomy.(26) We assumed that surveillance colonoscopies would be performed until at least 75 years of age, or until the stop age for screening, whichever was latest. If no adenomas were found at

diagnostic colonoscopy, the individuals were assumed to be at low-risk for CRC and did not return to the screening programme until after ten years.

Attendance

To identify the optimal screening strategies for participants, we analysed the strategies with full attendance (100%). In the sensitivity analysis, we looked at alternative gender-specific attendance levels based on the CORERO-1 trial, see Supplementary Table S1.

Costs

The analysis was performed from a third-party payer perspective. All costs are presented in Supplementary Table S1. We adjusted all costs to reflect the 2012 level, using the Dutch Consumer Price Index.(27)

FIT costs were assumed to be €21.90 based on an internal study (including a single KIT, packing material, material and personnel costs of the analysis, postage costs and organisational costs). The assumed costs of a colonoscopy were based on estimates in the COCOS trial: €192 for a negative colonoscopy and €329 for a colonoscopy with polypectomy.(28) Because of the recent discussion on colonoscopy costs in the US,(29) we considered costs that were twice and four times as high in a sensitivity analysis. Costs for colonoscopy complications were based on DTC-rates (Diagnosis Treatment Combination), derived from the Dutch Health Care Authority.(30) Costs for treatment of CRC were divided into three clinically relevant phases of care: initial treatment, continuous care and terminal care. The initial care phase was defined as the first 12 months after diagnosis; the terminal care phase was defined as the final 12 months of life; the continuing care phase was defined as all months in between. For patients surviving less than 24 months, the final 12 months were allocated to the terminal care phase, and the remaining months were allocated to the initial care phase. Initial treatment costs were based on DTC rates, except for oxaliplatin. The costs for oxaliplatin were derived from the Dutch Health Care Insurance Board.(31) We assumed that during the continuous care phase, individuals would follow the Dutch CRC treatment guidelines(32) and costs for periodic control were based on DTC rates. Terminal care costs were based on a Dutch last-year-of-life-cost-analysis.(33) We assumed that these costs increased with stage at diagnosis, at a rate observed for US patients.(34, 35) Dutch terminal care costs for individuals who died from CRC were approximately 40% of the US costs. We further assumed that terminal care costs of CRC patients who die from other causes were also 40% of the US costs.

Utility losses

We assumed no utility loss for a FIT, a utility loss equal to 2 days per colonoscopy (0.0055 QALYs) and two weeks of life per complication (0.0384 QALYs). We also assigned a utility loss to each life-year with CRC care (Supplementary Table S1).(36)

Analysis

Estimating FIT sensitivity and specificity

The sensitivity and specificity of FIT were fitted to the gender-specific positivity and detection rates observed in the first round of the CORERO-1 trial. FIT sensitivity and specificity were estimated by minimising the difference between observed and expected (i.e. model simulated) trial outcomes. Trial outcomes used for estimation were 1) positivity rate (PR), and detection rate (DR) of 2) CRC, 3) advanced adenomas and 4) non-advanced adenomas for both men and women, for a total of 8 trial outcomes. The observed detection rate of advanced adenomas was fitted to the detection rate of large (i.e. ≥10mm) adenomas in the model, since the model does not incorporate histology.

We estimated FIT characteristics twice: once assuming equal and once assuming gender-specific sensitivity and specificity. If the goodness-of-fit (GOF) of the model with gender-specific FIT characteristics was significantly better than the model with equal characteristics, we assumed that FIT characteristics indeed differed between men and women. The GOF was calculated as the sum of deviances between observed and simulated outcomes using the following formula:

$$2 * \left[obs * \left(ln\left(\frac{obs}{n}\right) - ln\left(\frac{sim}{m}\right)\right)\right] \; + \; 2 * \left[(n - obs) * \left(ln\left(\frac{n - obs}{n}\right) - ln\left(\frac{m - sim}{m}\right)\right)\right]$$

FIT characteristics differed significantly between men and women if the difference in GOF of the model with gender-specific FIT characteristics and the model assuming equal FIT characteristics exceeded 7.815 (chi-square distributed, three degrees of freedom).

Costs and benefits of uniform and gender-based screening

We used the MISCAN model to calculate costs and benefits of all 480 different screening strategies by gender, including no screening. Costs and QALYs gained were discounted by 3% per year.(37) For reference, we first compared outcomes of no screening, biennial screening from age 50-75 and annual screening from age 50-75 for men and women.

Subsequently, we used incremental cost-effectiveness analysis to determine the cost-effective screening strategies among all 480 evaluated screening by gender. To obtain these strategies, we ruled out strategies that were more costly and less effective

than other strategies (simple dominance) or combinations of other strategies (extended dominance). The remaining strategies are known as cost-effective or "efficient". On a plot of QALYs gained versus costs, the line connecting efficient strategies is called the efficient frontier. We calculated the incremental cost-effectiveness ratio (ICER) of each efficient strategy by comparing its costs and effects with those of the next less costly and less effective efficient strategy.

Finally, to determine the benefit of screening stratified by gender on a population level, we combined the efficient strategies of men with the efficient strategies of women, thereby creating the gender-based screening strategies. The costs and QALYs gained for men and women were summed, based on the distribution of men and women in the population. Then, the efficient gender-based screening strategies were determined and compared to the efficient strategies of uniform screening in the total population. We considered a difference in benefit between gender-based and uniform screening of ≥10% significant.(38)

Sensitivity analysis

We performed seven sensitivity analyses on different test characteristics of FIT: (i) we assumed only specificity differed between men and women; (ii) we assumed only sensitivity differed, (iii) we assumed no difference in sensitivity and specificity; (iv) we assumed a difference in sensitivity of CRC similar to the difference in sensitivity of advanced adenomas; (v) we assumed that sensitivity of FIT in women is primarily lower for progressive adenomas and to a lesser extent for non-progressive adenomas; (vi) we assumed that a percentage of adenomas do not bleed and can therefore never be detected by FIT, unless they grow and (vii) a similar analysis where we assumed that this percentage was higher in women than in men. (Supplementary Table S1)

We also performed sensitivity analyses on differential attendance for men and women, colonoscopy costs, treatment costs, discounting rates and including societal costs (Supplementary Table S1).

RESULTS

FIT characteristics

Assuming equal FIT characteristics for both sexes, the simulated PR and DR at a cut-off of 10 µg Hb/g faeces (50 ng Hb/ml) were higher in men than in women, due to a higher prevalence of colorectal neoplasia in men. Under this assumption, the simulated PR and DR in men were lower and in women higher than the observed rates (Table 1). Allowing FIT characteristics to vary by sex significantly improved the FIT of the model

to the observed CORERO-1 trial rates (Table 1). FIT specificity needed to be lower and sensitivity for (non)advanced adenomas needed to be higher in men than in women to replicate the observed FIT positivity and detection rates by gender, whereas the sensitivity for CRC was similar in both sexes (Supplementary Table S2).

The model simulated with equal FIT characteristics had a GOF of 56.3, compared to a GOF of 0.0008 in the model with genderspecific FIT characteristics, a difference of 56.3.

Table 1: **Positivity rates and detection rates FIT with a cut-off of 10 µg Hb/g faeces as simulated with equal and genderspecific test characteristics and as observed in CORERO-1.**

		Positivity rate	Detection rate of nonadvanced adenomas	Detection rate of advanced neoplasia*
MEN	Observed (N= 2857)	10.71%	2.56%	4.38%
	Simulated with *equal* FIT characteristics	8.60%	1.80%	3.48%
	Simulated with *genderspecific* FIT characteristics	10.75%	2.55%	4.38%
WOMEN	Observed (N= 3129)	6.30%	0.86%	2.17%
	Simulated with *equal* FIT characteristics	7.89%	1.50%	2.72%
	Simulated with *genderspecific* FIT characteristics	6.29%	0.86%	2.17%

*An advanced adenoma was defined as an adenoma of 10 mm or greater in size, and/or with 25% or greater villous component and/or high-grade dysplasia.

Screening outcomes by gender

Using the model with gender-specific test characteristics, MISCAN predicted that biennial FIT[10] screening between 50-75 years led to more profound reduction in CRC incidence and mortality compared to no screening in men than in women (Table 2). Women had less life years and QALYs gained than men per 1000 participants (35.7 versus 49.0 QALYs gained respectively), at higher costs (€42,161 versus -€5,471 respectively). Annual screening also yielded fewer QALYs gained (44.4 vs. 55.7) and higher costs (€68,555 vs. €15,391) in women than men when compared to no screening. However, the incremental QALYs gained and costs for annual screening compared to biennial screening were more similar between both sexes with 8.7 QALYs gained and €26,394 for women versus 6.7 QALYs gained and €20,863 for men.

When all strategies were considered (also varying screening age range and interval), costs remained higher and QALYs gained lower in women compared to men for all strategies (Figure 1). There was considerable overlap in which strategies were efficient between men and women, as six efficient screening strategies were identical (Table 3).

Table 2: **Outcomes of an annual and biennial screening program with FIT with a cut-off of 10 µg Hb/g feces and gender specific FIT characteristics, screening from age of 50-75 years per 1000 participants (100% attendance)**

		CRC incidence	CRC deaths	QALY's gained*	Total screening costs (€)*[ł]	Treatment costs (€)*	Total costs (€)*
	No screen	37.9	20.7		0	499,783	499,783
MEN	Biennial screening	23.3	9.6	49.0	136,267	358,045	494,312
	Annual screening	20.1	8.1	55.7	198,137	317,038	515,175
	No screen	31.9	18.5		0	420,600	420,600
WOMEN	Biennial screening	23.4	10.6	35.7	114,881	347,880	462,761
	Annual screening	19.8	8.6	44.4	182,660	306,495	489,155

* 3% discounted

[ł] including primary test, diagnostic colonoscopy, surveillance and complications

Benefit of gender-based screening

Supplementary Table S3 shows all efficient screening strategies stratified by gender and efficient uniform screening strategies. Six of these strategies included an identical screening strategy for men and women. Table 4 shows an example of uniform screening strategies and screening strategies stratified by gender at different willingness-to-pay thresholds. At a willingness-to-pay threshold of €0, screening stratified by gender consisted of screening men only. At a willingness-to-pay threshold of €20,000, the most effective strategy was equal in men and women, thus there was no difference between gender-stratified and uniform screening. The costs and QALYs gained of all efficient strategies are shown in Figure 2. For screening strategies with few screening rounds, screening stratified by gender dominated uniform screening, albeit the difference was small. The widest gap in QALYs gained between uniform screening and screening stratified by gender was at savings of €16,867: screening both men and women aged 60-70 years triennially gained less QALYs (54) than screening men aged 60-70 years biennially and screening women aged 60-70 years triennially (58) (Supplementary Table S3), a difference of 7%. From willingness-to-pay thresholds of €1,300 or higher, there was no difference between screening stratified by gender and uniform screening.

Sensitivity analysis

The performed sensitivity analyses resulted in different strategies to be on the efficient frontier. However, in all sensitivity analyses the added value of screening stratified by gender compared to uniform screening was marginal (Supplementary Figure S1). At a willingness-to-pay threshold of €20,000 per QALY gained there was no difference between uniform and stratified screening when assuming differential attendance levels for men and women. At this threshold level, the difference in QALY gained between uniform screening and screening stratified by gender was highest when assuming

quadruple colonoscopy costs, but did not exceed 3 QALY per 1000 participants (approximately 3%) (Panel B).

Table 3: **Screening strategies on the cost efficiency frontier in a) men and b) women, 3% discounted.**

	Cut-off	Start age	Stop age	Interval	# Screens	Costs*	QALY gained*	Costs (€) / QALY gained	ICER
					MEN				
1	FIT10	60	70	2	6	-54,815	76	-719	-719
2	FIT10	55	70	2	8	-48,971	87	-560	524
3	FIT10	55	70	1.5	11	-41,287	95	-434	998
4	FIT10	55	75	1.5	14	-32,956	100	-331	1,859
5	FIT10	50	75	1.5	17	-10,543	106	-99	3,525
6	FIT10	50	75	1	26	31,175	113	276	6,006
7	FIT10	50	80	1	31	63,867	116	551	10,809
8	FIT10	45	80	1	36	97,580	118	824	13,285
9	FIT10	40	80	1	41	125,815	119	1054	33,234
					WOMEN				
1	FIT10	60	70	3	4	18,948	41	462	462
2	FIT10	60	70	2	6	25,890	51	512	730
3	FIT10	60	70	1.5	7	28,696	54	533	860
4	FIT10	55	70	1.5	11	52,302	67	783	1,818
5	FIT10	55	75	1.5	14	66,794	72	921	2,549
6	FIT10	55	75	1	21	101,147	81	1241	3,823
7	FIT10	50	75	1	26	135,405	88	1543	5,449
8	FIT10	50	80	1	31	167,127	92	1812	7,131
9	FIT10	45	80	1	36	199,055	95	2092	10,818
10	FIT10	40	80	1	41	225,426	96	2341	23,267

*per 1000 participants;

Figure 1: **Costs and QALY's gained per 1000 participants of FIT screening with 5 different cut-offs and with different start and stop age and screening interval, 3% discounted.**

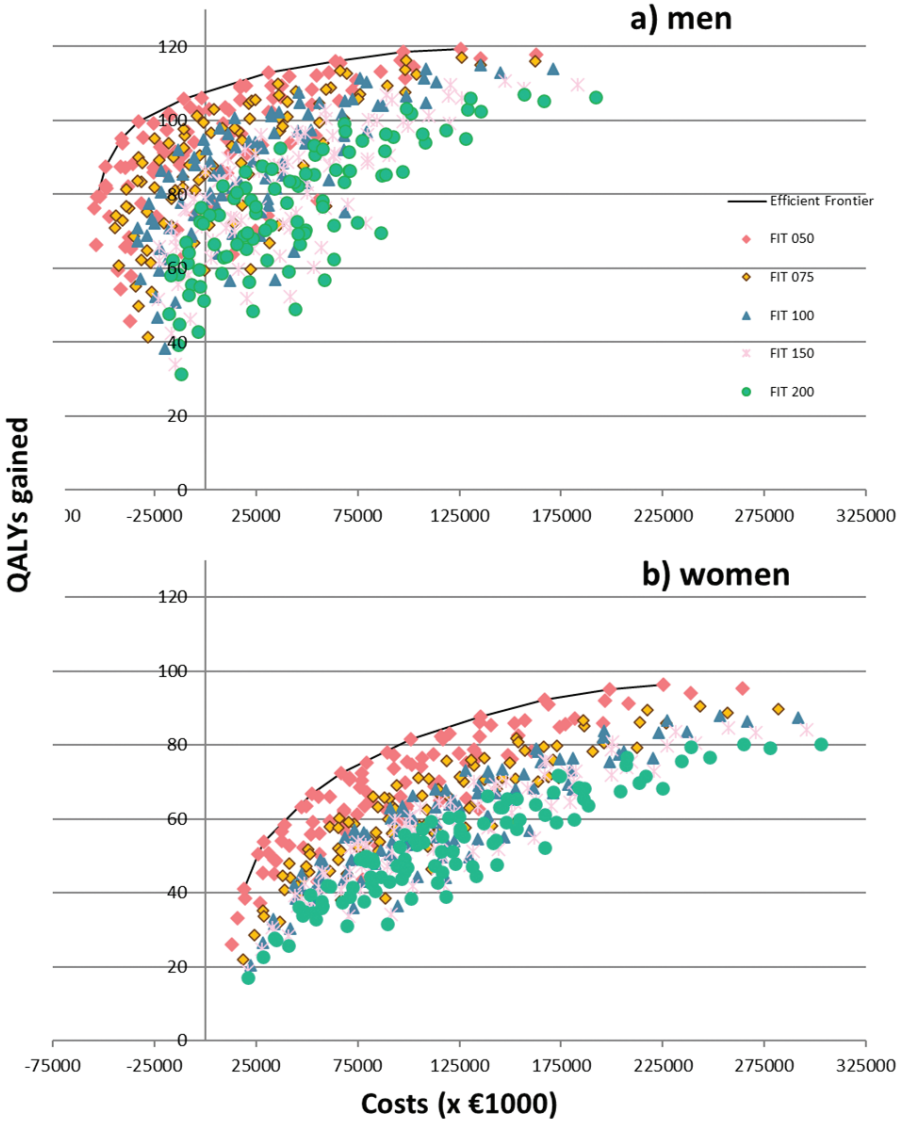
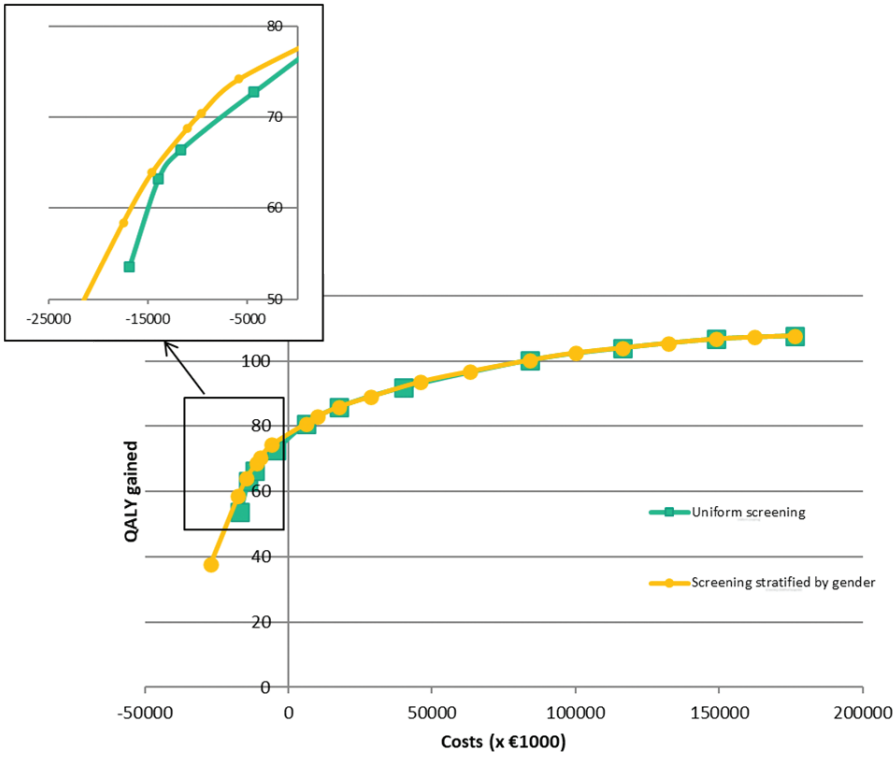


Figure 2: **Costs and QALY's gained of strategies on the cost efficiency frontier per 1000 participants with uniform screening and screening stratified by gender, 3% discounted.**



DISCUSSION

Our study demonstrates that FIT screening is more (cost-)effective in men than in women due to a higher prevalence of colorectal neoplasia and a better test sensitivity for (advanced) adenomas in men. Nevertheless, screening women remained highly cost-effective compared to no screening. Despite the difference in cost-effectiveness compared to no screening, the ICER of different screening strategies did not differ substantially between men and women and the optimal screening strategies for men and women were either the same or very similar. As a result, FIT screening stratified by gender dominated uniform screening with less intensive screening (maximum difference 58 versus 54 QALYs gained respectively), but resulted in equal costs and QALYs gained from a willingness-to-pay threshold of €1300. Thus, FIT screening stratified by gender was not more cost-effective than uniform FIT screening.

Given the differences in costs and QALYs gained compared to no screening between men and women, it may come as a surprise that the efficient strategies and incremental cost-effectiveness ratios are quite similar between sexes. Cost-effectiveness

of intensifying screening is however determined by the yield of the additional screening rounds. The prevalence of (advanced) neoplasia in the screened population decreases each screening round, depending on FIT sensitivity. At the first screening round, men have a higher prevalence of (advanced) neoplasia than women, but the prevalence of advanced neoplasia in men after one screening round will become lower than in unscreened women. As a consequence, the yield of initiating screening in women is higher than the yield of intensifying screening in men. This effect is demonstrated by a lower ICER of the first efficient strategy of women than the ICER of the second efficient strategy of men. The yield of further intensifying FIT screening depends on the residual number of non-detected neoplasia. The lower sensitivity of FIT in women compared to men necessitates more frequent screening in women than men, while the lower initial prevalence of neoplasia compensates this, leading to similar efficient strategies with a similar ICER.

We have modeled the differential performance of FIT between men and women as a difference in sensitivity for (advanced) adenomas. This does not necessarily mean that FIT is less accurate in women. Rather, adenomas in women are less likely to give blood in stool and therefore FIT is not able to detect these adenomas, resulting in a lower sensitivity of the test for adenomas. One explanation for the differential performance of FIT is the fact that a greater proportion of adenomas in men are generally located in the left hemicolon. Because this could influence results if the (missed) right sided lesions progress more rapidly, we added a sensitivity analysis in which the difference in FIT performance between men and women primarily existed for progressive (i.e. faster-growing) adenomas rather than for non-progressive (slow-growing) adenomas. The conclusion of this sensitivity analysis was in line with the base case analysis. Potential other reasons for the differential performance of FIT are gender differences in haemoglobin concentration of blood, faecal volume and a lower colonic transit in women than men.(5)

Our finding of lower test sensitivity of FIT in women is in concordance with two other studies, (6, 7) but in contrast with one other study.(39) Even though our sensitivity estimates are based on a single study, we are confident that this does not influence our results. We performed extensive sensitivity analyses on test characteristics and found our results to be robust for these assumptions. Also, a German study found a per person sensitivity for advanced neoplasia of 30.7% for women compared to 47.7% for men.(6) Our sensitivity estimates concern a per-lesion instead of per-person sensitivity. Due to multiple lesions and a probability for a positive FIT for other reasons than colorectal neoplasia (e.g. hemorrhoids), the sensitivity on a person-level is higher than the per-lesion sensitivity. Our per-person sensitivity as calculated from the model output is quite similar to the German study (32.5% for women, versus 55.4% for men).

Our results are obtained by assuming perfect (100%) attendance, because assuming imperfect adherence could result in overly aggressive screening in hope that, on average, screening is performed at the desired frequency. This would lead to overscreening in those who adhere with recommendations, with the potential for unnecessary risks. However, we showed in the sensitivity analyses that assuming gender-specific realistic attendance did not influence our conclusions.

To our knowledge, this is the first cost-effectiveness analyses to determine the optimal FIT screening strategy by gender. Two limitations are noteworthy. First, we assumed that all differences in the prevalence of adenomas and CRC incidence between men and women were caused by a difference in adenoma onset and probability to progress to CRC. We did not assume any differences in dwelling time of adenomas. However, since the relative risk for men and women of non-advanced adenomas in a German study is similar to the relative risks of CRC in the Netherlands in the corresponding age group (RR 1.5),(6) we believe it is likely that the dwelling time of adenomas does not differ significantly between men and women. Second, we introduced a sensitivity analysis in which a proportion of adenomas are systematically missed, but assumed this proportion was equal for men and women. If this proportion does differ, it might influence the preferred screening ages and interval, in theory making screening stratified by gender more beneficial. There are not enough data yet to study this phenomenon for men and women separately, but we did include a sensitivity analysis with a hypothetical difference in the proportion, showing the same conclusion in this sensitivity analysis as the base case analysis.

Various investigators have argued that CRC screening should be stratified by gender because of the difference in prevalence of (advanced) neoplasia(40, 41) and the gender-related differences in FIT accuracy.(5, 9) Our study shows that the added value of gender-based screening is at most marginal. Furthermore, screening stratified by gender may also have disadvantages: some men and women may be confused by the differential recommendations to the point that they no longer attend screening. A slight impact of stratified screening recommendations on attendance will easily offset its marginal benefit. On the other hand, screening stratified by gender may increase attendance because participants feel that the recommendations are better tailored to their risk. Therefore, future research is needed in this area.

Another area for future research is to evaluate the comparative effectiveness of FIT screening and other screening modalities in men and women separately. Earlier studies showed not much difference in cost-effectiveness between a FIT screening programme and colonoscopy screening for the population as a whole.(42) However, since sensitivity of FIT is lower in women than men, the additional sensitivity of colonoscopy compared to FIT is also higher in women leading to lower comparative effectiveness of FIT with

colonoscopy. If the lower sensitivity of FIT in women does not apply to other stool-based tests, the comparative effectiveness of newer tests, such as stool-DNA tests, could also be different than in men.

In conclusion, this study shows that the (cost)effectiveness of FIT screening is higher in men than in women due to a higher FIT sensitivity and a higher prevalence of neoplasia in men. However, optimal screening strategies were similar in men and women with respect to interval, age range and FIT cut-off. Screening stratified by gender does not improve cost-effectiveness and therefore our findings support uniform screening of men and women as currently applied in FIT screening programmes, like in the Netherlands.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61:69-90.
2. Gupta S, Sussman DA, Doubeni CA, Anderson DS, Day L, Deshpande AR, et al. Challenges and possible solutions to colorectal cancer screening for the underserved. J Natl Cancer Inst. 2014;106:dju032.
3. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64:1637-49.
4. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Winawer SJ, et al. Individualizing colonoscopy screening by sex and race. Gastrointest Endosc. 2009;70:96-108, e1-24.
5. Steele RJ, McClements P, Watling C, Libby G, Weller D, Brewster DH, et al. Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. Gut. 2012;61:576-81.
6. Brenner H, Haug U, Hundt S. Sex differences in performance of fecal occult blood testing. Am J Gastroenterol. 2010;105:2457-64.
7. Stegeman I, de Wijkerslooth TR, Stoop EM, van Leerdam M, van Ballegooijen M, Kraaijenhagen RA, et al. Risk factors for false positive and for false negative test results in screening with fecal occult blood testing. Int J Cancer. 2013;133:2408-14.
8. Kapidzic A, van der Meulen MP, Hol L, van Roon AH, Looman CW, Lansdorp-Vogelaar I, et al. Gender Differences in Fecal Immunochemical Test Performance for Early Detection of Colorectal Neoplasia. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015.
9. Isabel Portillo CF. Should FIT cut-offs be gender-specific? 2014 29/10/2015]; Available from: http://www.worldendo.org/assets/downloads/pdf/resources/ccsc/2014/ueg/06_CGF_and_IPV_for_CRCSC_Vienna.pdf
10. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JC, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59:62-8.
11. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009;100:1103-10.
12. Lansdorp-Vogelaar I, van Ballegooijen M, Boer R, Zauber A, Habbema JD. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. Cancer. 2009;115:2410-9.
13. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res. 1999;32:13-33.
14. Comprehensive Cancer Centre. [cited 2012 14 Mar]; Available from: http://cijfersoverkanker.nl/
15. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. Cancer. 1988;61:1472-6.
16. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. Ann Surg. 1963;157:223-6.
17. Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. Int J Cancer. 1985;36:179-86.

18. Jass JR, Young PJ, Robinson EM. Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. Gut. 1992;33:1508-14.
19. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. Scand J Gastroenterol. 1989;24:799-806.
20. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. Cancer. 1979;43:1847-57.
21. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. Cancer. 1982;49:819-25.
22. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. Gut. 1982;23:835-42.
23. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, 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;13:55-64.
24. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375:1624-33.
25. Statistics Netherlands. [cited 2012 14 Mar]; Available from: http://statline.cbs.nl/StatWeb/default.aspx
26. CBO guideline. [cited 2014 25 Sep]; Available from: http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf
27. Statistics Netherlands. [cited 2012 14 March]; Available from: http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=70936ned&D1=a&D2=584,597,610,623,636,649&STB=T,G1&VW=T
28. Stoop EM. Population-based colorectal cancer screening by colonoscopy or CT-colonography [thesis]. Rotterdam: Erasmus University Rotterdam; 2013.
29. 2013 [cited 2014 9 Oct]; Available from: http://www.nytimes.com/2013/06/02/health/colonoscopies-explain-why-us-leads-the-world-in-health-expenditures.html?pagewanted=all&_r=1&
30. Dutch Health Care Authority. [website]. [cited 2012 14 Mar]; Available from: http://www.nza.nl/regelgeving/tarieven/
31. College voor zorgverzekeringen (CVZ); Farmacotherapeutisch Kompas 2014.
32. Guidelines for Oncologic Treatments. [cited 2012 14 Mar]; Available from: http://www.oncoline.nl
33. de Kok IM, Polder JJ, Habbema JD, Berkers LM, Meerding WJ, Rebolj M, et al. The impact of healthcare costs in the last year of life and in all life years gained on the cost-effectiveness of cancer screening. Br J Cancer. 2009;100:1240-4.
34. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Habbema JD, Kuipers EJ. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. J Natl Cancer Inst. 2009;101:1412-22.
35. Yabroff KR, Lamont EB, Mariotto A, Warren JL, Topor M, Meekins A, et al. Cost of care for elderly cancer patients in the United States. J Natl Cancer Inst. 2008;100:630-41.
36. Ness RM, Holmes AM, Klein R, Dittus R. Utility valuations for outcome states of colorectal cancer. Am J Gastroenterol. 1999;94:1650-7.

37. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. Pharmacoeconomics. 1997;11:159-68.
38. Donken R, de Melker HE, Rots NY, Berbers G, Knol MJ. Comparing vaccines: a systematic review of the use of the non-inferiority margin in vaccine trials. Vaccine. 2015;33:1426-32.
39. Wong MC, Ching JY, Chan VC, Lam TY, Luk AK, Ng SS, et al. Factors associated with false-positive and false-negative fecal immunochemical test results for colorectal cancer screening. Gastrointest Endosc. 2015;81:596-607.
40. Lieberman D. Race, gender, and colorectal cancer screening. Am J Gastroenterol. 2005;100:2756-8.
41. Regula J, Kaminski MF. Targeting risk groups for screening. Best Pract Res Clin Gastroenterol. 2010;24:407-16.
42. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev. 2011;33:88-100.

SUPPLEMENTARY MATERIAL

Supplementary Table S1: **Summary of model assumptions of the base case and sensitivity analyses**

Variable	Base Analysis	Sensitivity analysis
Test characteristics FIT		
	Calibrated to CORERO-1 data separately for men and women, see Appendix 2	Equal test characteristics, see Appendix 3
		Only difference in sensitivity/ Only difference in specificity / Difference in CRC sensitivity with same rate as difference in advanced adenomas
	No systematic FIT failure	Proportion of adenomas with systematic FIT failure (non-bleeding adenomas)
Sensitivity colonoscopy		
1-5 mm adenomas	75%	
6-9 mm adenomas	85%	
≥10 mm adenomas	95%	
carcinomas	95%	
Adherence		
Screening test	100%	59.8% for men and /64.6% for women
Diagnostic test	100%	97.1% for men and /95.6% for women
Surveillance test	100%	80%
Quality of life loss		
Colonoscopy	2 days lost per colonoscopy	
CRC from diagnosis onward	Initial treatment stage 1 till IV: 0.12; 0.18; 0.24; 0.70	
	Continuous care stage 1 till IV: 0.12; 0.18; 0.24; 0.70 Terminal care death by CRC: 0.70	
	Terminal care death by other cause: 0.12; 0.18; 0.24; 0.70	
Fatal complications after colonoscopy	3.29^10^{5} in positive colonoscopies	
Screening costs		
FIT	€ 21.90	
Diagnostic costs inside screening program (positive/negative)	€329/€192	Double colonoscopy costs / quadruple colonoscopy costs
Costs complications after colonoscopy	€ 1,372	
Treatment costs*		
	Initial Continuous Terminal care Terminal care	
	care care death CRC death other c cause	

Supplementary Table S1: **Summary of model assumptions of the base case and sensitivity analyses**

Variable	Base Analysis	Sensitivity analysis
stage I	€13,773 €375 €19,282 €4,848	Half / Double
stage II	€18,180 €375 €19,282 €4,407	
stage III	€20,935 €375 €20,384 €5,729	
stage IV	€27,546 €375 €27,546 €15,426	
Discounting	3%	No discounting/ 1.5% discounting on QALY's and 4% on costs
Perspective	Third party payer	Societal costs[†]

***** Care for CRC was divided in 3 clinically relevant phases: the initial, continuing, and terminal care phases. The initial care phase was defined as the first 12 months after diagnosis; the terminal care phase was defined as the final 12 months of life; the continuing care phase was defined as all months in between. In the terminal care phase, we distinguished between CRC patients dying from CRC and CRC patients dying from another cause. For patients surviving less than 24 months, the final 12 months were allocated to the terminal care phase, and the remaining months were allocated to the initial care phase. [†] Patient time costs were added (ie, the opportunity costs of spending time on screening or being treated for a complication or CRC) but we did not include travel costs, costs of lost productivity, and unrelated health care and non-health care costs in added years of life. We assumed that the value of patient time was equal to the median wage rate in 2012: €15.93 per hour (Statistics Netherlands) We assumed that colonoscopies and complications used 8 and 16 hours of patient time, respectively. Patient time costs for treatment of CRC were based on a study by Yabroff et al. (K.R. Yabroff, E.B. Lamont, A. Mariotto, **et al.**; Cost of care for elderly cancer patients in the United States; J Natl Cancer Inst, 100 (2008), pp. 630–641) and corrected for the Dutch median wage rate.

Supplementary Table S2: **Calibrated specificity and sensitivity per adenoma of the FIT to CORERO-1 data for men and women.**

	Men	Women	Total population
FIT 10 µg Hb/g feces			
Specificity	95.0%	95.9%	95.5%
Sensitivity per nonadvanced adenoma	19.1%	1.0%*	10.0%
Sensitivity per advanced adenoma	46.7%	26.5%	34.3%
Sensitivity per crc long before clinical diagnosis	46.7%†	42.9%	45.0%
Sensitivity per crc short before clinical diagnosis	78.4%	77.8%	79.2%
FIT 15 µg Hb/g feces			
Specificity	97.1%	97.1%	97.1%
Sensitivity per nonadvanced adenoma	14.2%	1.0%*	7.1%
Sensitivity per advanced adenoma	42.1%	20.3%	29.7%
Sensitivity per crc long before clinical diagnosis	43.2%‡	37.8%‡	41.1%
Sensitivity per crc short before clinical diagnosis	78.0%‡	73.8%‡	76.4%
FIT 20 µg Hb/g feces			
Specificity	97.7%	97.8%	97.8%
Sensitivity per nonadvanced adenoma	11.3%	1.0%*	6.1%
Sensitivity per advanced adenoma	38.1%	18.1%	26.8%
Sensitivity per crc long before clinical diagnosis	43.2%	37.8%	42.4%
Sensitivity per crc short before clinical diagnosis	78.0%	73.8%	77.4%
FIT 30 µg Hb/g feces			
Specificity	98.5%	98.3%	98.4%
Sensitivity per nonadvanced adenoma	9.0%	1.0%*	4.9%
Sensitivity per advanced adenoma	33.1%	17.3%	24.2%
Sensitivity per crc long before clinical diagnosis	35.4%	32.4%	35.0%
Sensitivity crc short before clinical diagnosis	71.8%	69.0%	71.4%
FIT 40 µg Hb/g feces			
Specificity	98.7%	98.7%	98.7%
Sensitivity per nonadvanced adenoma	7.8%	1.0%*	4.5%
Sensitivity per advanced adenoma	30.0%	15.6%	21.9%
Sensitivity per crc long before clinical diagnosis	30.0%†	27.0%	28.5%
Sensitivity per crc short before clinical diagnosis	63.8%	63.2%	65.0%

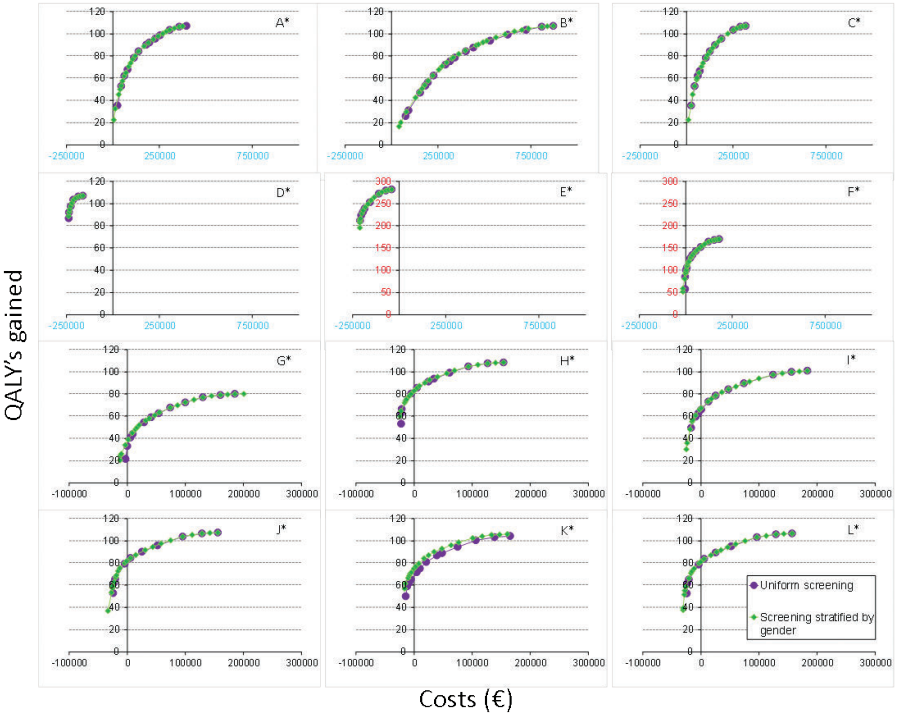
* Sensitivity per nonadvanced adenoma in women varied slightly over the cut-offs but did not decrease with higher cut-off, we decided to use the average sensitivity for each cut-off.
† Sensitivity per colorectal carcinoma long before clinical diagnosis was lower then of advanced adenomas at the same cut-off, therefore we assumed the same sensitivity as for advanced adenomas.
‡ Sensitivity per colorectal carcinoma was lower then of colorectal carcinoma at a higher cut-off, therefore we assumed the same sensitivity as the higher cut-off.

Supplementary Table S3: **Screening strategies on the cost efficiency frontier with uniform screening and all combined strategies for screening stratified by gender, 3% discounted.**

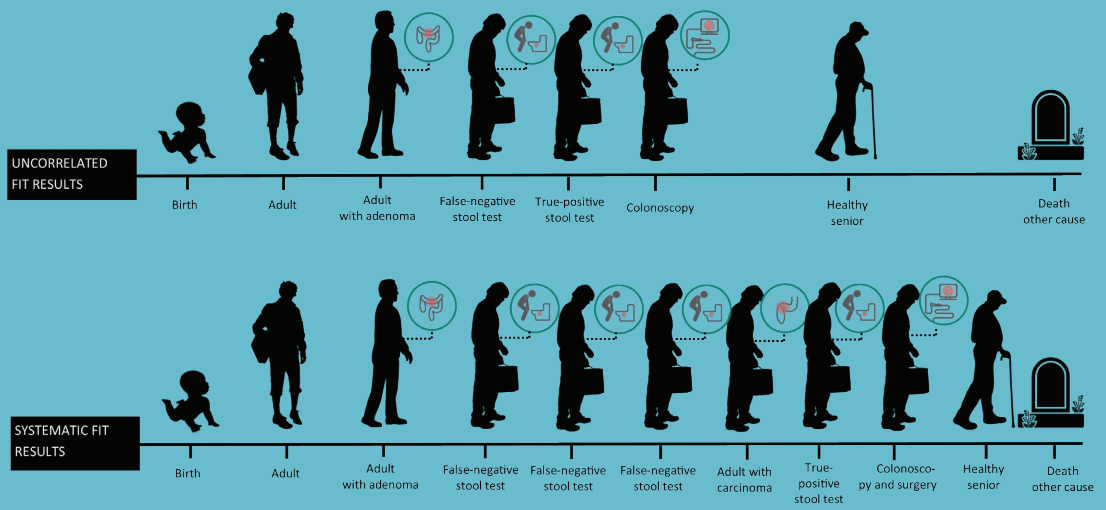
Cut-off	Start age		Stop age		Interval		# Screens		Costs*	QALY gained*	Costs/ QALY gained	ICER
	M	W	M	W	M	W	M	W				
DIFFERENTIAL SCREENING												
FIT10	60	x	70	x	2	x	6	x	-27,062	38	-719	-719
FIT10	60		70		2	3	6	4	-17,469	58	-299	462
FIT10	55	60	70		2	3	8	4	-14,584	64	-228	524
FIT10	55	60	70		2		8	6	-11,069	69	-161	730
FIT10	55	60	70		2	1.5	8	7	-9,649	70	-137	860
FIT10	55	60	70		1.5		11	7	-5,855	74	-79	998
FIT10	55		70		1.5		11		6,096	81	75	1,818
FIT10	55		75	70	1.5		14	11	10,209	83	123	1,859
FIT10	55		75		1.5		14		17,547	86	204	2,549
FIT10	50	55	75		1.5		17	14	28,612	89	321	3,525
FIT10	50	55	75		1.5	1	17	21	46,005	84	492	3,823
FIT10	50		75		1.5	1	17	26	63,350	97	655	5,449
FIT10	50		75		1		26		83,946	100	838	6,006
FIT10	50		75	80	1		26	31	100,007	102	976	7,131
FIT10	50		80		1		31		116,147	104	1,118	10,809
FIT10	50	45	80		1		31	36	132,312	105	1,255	10,818
FIT10	45		80		1		36		148,956	107	1,396	13,285
FIT10	45	40	80		1		36	41	162,308	107	1,513	23,267
FIT10	40		80		1		41		176,247	108	1,637	33,234
UNIFORM SCREENING												
FIT10	60		70		3		4		-16,867	54	-315	-315
FIT10	60		70		2		6		-13,954	63	-221	302
FIT10	60		70		1.5		7		-11,725	66	-177	706
FIT10	55		70		2		8		-4,393	73	-60	1,155
FIT10	55		70		1.5		11		6,096	81	75	1,307
FIT10	55		75		1.5		14		17,547	86	204	2,249
FIT10	50		75		1.5		17		40,020	92	436	3,752
FIT10	50		75		1		26		83,946	100	838	5,285
FIT10	50		80		1		31		116,147	104	1,118	8,597
FIT10	45		80		1		36		148,956	107	1,396	11,943
FIT10	40		80		1		41		176,247	108	1,637	27,476

*per 1000 participants

Supplementary Figure S1: **Costs and QALY's gained of strategies on the cost efficiency frontier per 1000 participants with uniform screening and screening stratified by gender with different assumptions in the sensitivity analysis, 3% discounted if not other specified**



Sensitivity analysis with: A, double colonoscopy costs; B, quadruple colonoscopy costs; C, half treatment costs; D, double treatment costs; E, 0% discounting; F, QALY's gained 1.5% discounted and costs 4% discounted; G, societal costs; H, differential attendance as observed in the CORERO-trial; I, equal FIT characteristics for men and women; J, differing sensitivity in men and women but equal specificity; K, differing specificity in men and women but equal sensitivity; systematic FIT failure L, equal for men and women; M, with a hypothetical difference; N, Difference in sensitivity between men and women primarily in progressive adenomas; O, differing CRC sensitivity additional to other sensitivity and specificity difference


UNCORRELATED FIT RESULTS
Birth
Adult
Adult with adenoma
False-negative stool test
True-positive stool test
Colonoscopy
Healthy senior
Death other cause
SYSTEMATIC FIT RESULTS
Birth
Adult
Adult with adenoma
False-negative stool test
False-negative stool test
False-negative stool test
Adult with carcinoma
True-positive stool test
Colonosco-py and surgery
Healthy senior
Death other cause

Chapter 4

Non-bleeding adenomas: evidence of systematic false-negative fecal immunochemical test results and its implications for screening effectiveness – a modeling study.

Miriam P. van der Meulen, Iris Lansdorp-Vogelaar, Else-Mariëtte B. van Heijningen, Ernst J. Kuipers, Marjolein van Ballegooijen

Cancer. 2016;122(11):1680-8.

ABSTRACT

Background: If some adenomas do not bleed over several years, they will cause systematic false-negative fecal immunochemical test (FIT) results. Long-term effectiveness of FIT screening has been estimated without accounting for such systematic false-negativity. There are now data to evaluate this issue.

Methods: We developed one micro-simulation model MISCAN-Colon without systematic false-negative FIT results and one that allowed a proportion of adenomas to be systematically missed in successive FIT screening rounds. Both variants were adjusted to reproduce the first-round findings of the Dutch CORERO FIT screening trial. We then compared simulated detection rates in the second screening round with those observed, and adjusted the simulated proportion of systematically missed adenomas to those data. Finally, we calculated the impact of systematic false-negative FIT results on the effectiveness of repeated FIT screening.

Results: The model without systematic false-negativity simulated higher detection rates in the second screening round than observed. These observed rates could be reproduced when assuming that FIT systematically missed 26% of advanced and 73% of non-advanced adenomas. To reduce the false positive rate in the second round to the observed level, we also had to assume that 30% of false-positives were systematically false-positive. Systematic false-negative FIT testing limits the long-term reduction of biennial FIT screening in CRC incidence (35.6% versus 40.9%) and mortality (55.2% versus 59.0%) in participants.

Conclusion: This study provides convincing evidence based on the combination of real-life and modeling data that a proportion of adenomas is systematically missed by repeated FIT screening. This impairs the efficacy of FIT screening.

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. Repeated screening by means of the guaiac fecal occult blood test (gFOBT) reduces mortality by 15%-33% as shown in several trials.[2-4] Since these trials, more sensitive FOBTs have been developed. One new version is the fecal immunochemical test (FIT), which is specific for the detection of human globin in stool. Several trials showed that FIT screening is associated with a higher diagnostic yield and higher adherence than gFOBT screening.[5-11] As a consequence, several countries, such as Italy, Australia, Japan and The Netherlands, have adopted population-based FIT screening,[12] while for instance in the United States several local initiatives recently adopted FIT screening.[12,13]

New tests always raise the question whether randomized controlled trials, which are expensive and take at least ten years before producing results, are necessary. The long-term effectiveness of repeated FIT screening is not yet studied empirically. Meanwhile, modeling studies have been used to extrapolate the higher diagnostic yield of FIT compared to gFOBT to determine the long-term effectiveness of FIT screening. These studies all assumed that FOBT results over consecutive screening rounds are independent of each other,[14] also implying that the increase in sensitivity when moving from gFOBT to FIT holds over the rounds.

However, little is known about when adenomas or carcinomas start to bleed and how often they bleed.[15,16] The gFOBT trials already gave an indirect measure of bleeding patterns in carcinomas,[17] but data on adenomas were scarce. If adenomas of similar size do not have an equal chance to bleed, they also do not have an equal chance to be picked up by FIT. This would imply that FIT results over subsequent rounds depend on each other: previous false negative results increase the probability of another such result. As long as adenomas do not bleed at all, they will cause persistent, so-called systematic false-negative FIT results. Without accounting for this phenomenon, one might overestimate the effectiveness of a screening program with FIT.[18]

Data from repeat screening show that the diagnostic yield of adenomas decreases with consecutive screening rounds.[19-21] This is expected, due to depletion of the prevalence pool of adenomas. However, if systematic false negative test results occur, the diagnostic yield will decrease even further than explained by that depletion.

We used the micro-simulation model MISCAN-Colon to compare simulated detection rates of repeated FIT screening with those observed in a Dutch population-based screening study. We studied to what extent incorporating systematic false-negative FIT results is necessary to explain those observed rates. Finally, we calculated

the impact of the estimated amount of systematic false-negative FIT results on the reduction in incidence and mortality of a FIT screening program.

MATERIALS AND METHODS

Overview

We used the MISCAN-Colon model to reproduce the design and first-round findings of the Dutch CORERO FIT screening trial.[7,20,22] We then compared detection rates in the second screening round simulated by the model without systematic false-negative FIT results with those observed in the trial. Subsequently, we developed a variant of the model with systematic false-negative FIT results and estimated how much systematic false-negative FIT results were needed to optimally fit the adenoma detection rates. We also compared the observed second round false-positive rate with the simulated rate to account for the possible effect of systematic false-positive FIT results. The combination of systematic false-positive and false-negative FIT results was defined as "systematic FIT failure". We then validated the estimates for systematic FIT failure with the third round of the CORERO trial and with first-round observations of the CORERO trial in the group that performed a 2-sample FIT. Finally, we calculated the impact of systematic FIT failure on the effectiveness of FIT screening by running screening programs by the models with and without systematic FIT failure.

The CORERO trial

The CORERO-phase-I trial was a randomized controlled trial comparing attendance and yield of gFOBT with those of FIT and sigmoidoscopy at first round screening.[7] Subsequently, CORERO-phase-II looked at the attendance and detection rates of repeat FIT screening at different intervals.[20] Details from these trial-phases have been described elsewhere.[7,20,22,23] In short, screening-naive subjects between 50-74 years old, living in the southwest of the Netherlands were selected through municipal population registers. In the first round, people were sent a kit with a single FIT test (OC-Sensor, Eiken, Japan). A cut-off of 10 µg hemoglobin/g feces (equivalent to 50 ng hemoglobin/mL) was used to indicate a positive test result after which a colonoscopy was offered by experienced endoscopists. A second screening round took place after an interval of one, two or three years (group I, II and III respectively).[20] A third screening round took place after an interval of two years for all three groups.[23] In the first screening round 4523 (62.6%) returned the test kit. Of these participants, 3427 (90.6%) responded in the second screening round. Because we were interested in systematic FIT failure in

the present analysis, we only considered participants of both screening rounds for the main analysis and for all three rounds for the validation to the third round.

CORERO-phase-I also contained a group which received a kit with 2 FIT samples (group IV).[22] They were instructed to conduct the 2 tests on subsequent days. Of this group, 1876 (61.2%) invitees returned the test kit.

MISCAN-Colon

MISCAN-Colon is a microsimulation model for CRC developed at the Department of Public Health of the Erasmus University Medical Center (Rotterdam, the Netherlands). The models structure, underlying assumptions, and calibration are described in the Model Appendix and previous publications.[24,25] In brief, the MISCAN-colon model simulates the relevant life histories of a large population of individuals from birth to death. CRC arises in this population according to the adenoma-carcinoma sequence.[26,27] More than one adenoma can occur in an individual and each adenoma can independently develop into a CRC. Adenomas may progress in size from small (≤5 mm) to medium (6-9 mm) to large (≥10 mm). Although most adenomas will never turn into cancer, some will eventually become malignant. Cancer starts as a symptomless process and can progress from localized cancer stage I to metastasized cancer stage IV. In every stage, there is a probability of the CRC being diagnosed due to the development of symptoms versus symptomless progressing into the next stage. At any time during the development of the disease, the process may be interrupted because a person dies of other causes. With FIT screening lesions can be detected before clinical diagnosis; a screened individual with a positive test result will be referred for a colonoscopy for the detection and removal of adenomas and cancers, hopefully in an earlier stage. In this way, CRC incidence and/or CRC-related mortality can be reduced.

MISCAN-Colon was calibrated to the age-, stage-, and localization-specific incidence of CRC as observed in the Netherlands before the introduction of screening (i.e., between 1999 and 2003)[28] and the age-specific prevalence and multiplicity distribution of adenomas as observed in autopsy studies.[29-37] The size distribution of adenomas was calibrated to the size distribution detected with colonoscopy in the COCOS-trial.[38] The preclinical duration of CRC and the adenoma dwell-time were calibrated to the rates of interval and surveillance-detected cancers observed in randomized controlled trials evaluating screening using guaiac fecal occult blood tests,[2-4] and a once-only sigmoidoscopy[39] and showed good concordance with the mortality reduction observed (Appendix 1).

FIT characteristics

Test characteristics of the one-sample FIT were fitted to the positivity and detection rates of non-advanced and advanced adenomas and carcinomas as observed in the first screening round of the CORERO-trial and its counterpart in two other regions of the Netherlands.[8,10]

Sensitivity and systematic false-negative FIT results

An advanced adenoma was defined as an adenoma ≥10 mm in size, with 25% or greater villous component and/or high-grade dysplasia. Because the model does not incorporate histology, the observed detection rate of non-advanced and of advanced adenomas was fitted to the detection rate of small to medium and of large adenomas in the model respectively. We modelled sensitivity by giving each lesion a probability to cause a positive FIT. The fitted per lesion probabilities were 0% for adenomas 1-5 mm, 11.4% for adenomas 6-9 mm, 34.4% for adenomas ≥ 10 mm, 50.3% for carcinomas long before clinical diagnosis and 82.5% for carcinomas shortly before clinical diagnosis.

We then included systematic false-negative FIT results by simulating the following concept. When an adenoma starts to develop, it will not bleed at first. During that phase, it will not have any chance to cause a positive FOBT. Once it starts to bleed, it has a random chance, based on the sensitivity, to cause a positive FOBT. We simulated this process in the model by discerning adenomas that already bleed with a random sensitivity, and adenomas that do not bleed yet with zero sensitivity, a systematic false-negativity. The random sensitivity was corrected upwards, such that the overall sensitivity for adenomas in the first screening round was not affected.

The probability of a systematic false-negative FIT result was estimated separately for advanced and non-advanced adenomas.

Note that all types of sensitivity are set per lesion. If a person had a second adenoma or a colorectal carcinoma that generated a positive test result, or had a positive test result due to the lack of specificity (chance detection), an adenoma that was “missed” (it did not by itself generate a positive test) could still be detected through diagnostic colonoscopy.

Specificity and systematic false-positive FIT results

In previous modeling studies we modeled lack of specificity as a per person probability of having a positive test, independent of whether this person had neoplasia. The fitted probability in this study was 95.7%.

The same concept as systematic false-negative FIT results can in principle also occur in false-positivity. This would e.g. occur if individuals have a constitution or (chronic) condition causing fecal bleeding over a number of years. In case individuals,

who had a negative colonoscopy after a positive FIT, are put on hold for several FIT screening rounds (as was done in the CORERO-trial), this results in a depletion of the prevalence pool of such individuals and thus in lower than expected false positive rates in later rounds.

We modelled systematic false-positive test results by assuming that a proportion of people would always test false-positive. We did this by assigning these individuals with a (0% specificity) systematic (false-)positivity, as opposed to individuals with a random specificity. The random specificity was increased so that the introduction of systematic false-positivity did not affect the overall number of false-positive test results in a first screening round.

Colonoscopy characteristics

The sensitivity of diagnostic colonoscopies was assumed to be 75% for adenomas ≤ 5 mm, 85% for adenomas 6-9 mm, and 95% for adenomas ≥10 mm and CRC.[40]

Data analysis

Calibration of systematic FIT failure

At first, a population was simulated with birth years that matched those of the invitees of the first screening round of the CORERO-trial. This was done separately for each group (group I till IV). The FIT screening strategies (one day interval for the 2-sample group, and 1, 2 or 3 year interval for the 1-sample groups), the attendance over the two screening rounds and the compliance to diagnostic colonoscopy after a positive FIT were also matched to the observed data. For both the observed and simulated data, we present aggregated data for the three groups in this study.

Then, we simulated two consecutive screening rounds with a 1-sample FIT at a cut-off of 10 microgram Hb/g feces in 10 million people with the model without systematic FIT failure. The positivity and detection rates and the positive predictive value (PPV) of the second round were determined and compared with the observed rates and their 95% confidence intervals in the CORERO-trial. The positivity and detection rates were determined by dividing the number of events (persons with a positive FIT result, persons with a detected adenoma) by the number of people screened.

Subsequently, the size of the systematic component of specificity and sensitivity was estimated by minimizing the difference between the observed and simulated rates in the second screening round. We first estimated the size of the systematic component of specificity on the observed difference in false positive rate between the first and second screening. We then estimated the probability of systematic false-negative test results for non-advanced adenomas and for advanced adenomas on the observed second round detection rate of non-advanced adenomas and advanced adenomas.

Validation of systematic FIT failure

To validate the systematic model, we used both the version with and without systematic FIT failure of the MISCAN-Colon model to project positivity and detection rates in subjects undergoing a third screening round and compared these to the observed third screening round of the CORERO trial. We also simulated 2-sample FIT results with and without systematic FIT failure and compared these to the 2-sample observed data of group IV in the CORERO-trial, to validate the model in a different dataset.

Effectiveness

We simulated a Dutch population born in 1955 until death that all attended the screening protocol as was introduced in the Netherlands in 2014: with a starting age of 55, a stopping age of 75 and a 2-year interval. We considered full attendance to explore the effect on individuals who comply with the complete program. First, we assumed no systematic FIT failure and then we introduced systematic false-negativity, systematic false-positivity and finally, both. We compared the mortality reduction, incidence reduction, lifetime number of colonoscopies per person, number needed to scope to prevent one death (NNScope), life years gained and quality adjusted life years (QALYs) gained of all the four scenarios. The life years and QALYs gained were undiscounted and alternatively discounted with 3%.

In addition, to show the maximum possible effect of systematic FIT failure, we repeated these runs assuming that all first-round false-negative adenomas were non-bleeding adenomas.

RESULTS

The CORERO-trial

In the first round of the CORERO trial 8.4% of the participants had a positive FIT, in the second round the positivity rate went down to 5.8%.[20] This decline was caused by a decline in both false positive test results (false positive rate) and true positive test results. The decline in true positive test results was reflected in the detection rate of non-advanced adenomas, advanced adenomas and colorectal carcinomas, see Table 1.

Simulated results of the first round

The model successfully reproduced the observed positivity and detection rates of the FIT of the first round. As explained in the methods, the introduction of systematic FIT failure did not affect the simulated positivity and detection rates of the first screening round, see Table 1.

Simulated results of the second round

In the second round, the model without systematic FIT failure had a considerably higher positivity rate than observed. While the simulated positivity rate declined 1.0 percentage point between first and second round (from 8.3% to 7.3%), the observed decline was 2.6 percentage points (from 8.4% to 5.8%). This smaller simulated decline resulted both from a smaller decline in false positive rate and a smaller decline in simulated detection rates. The false positive rate simulated by the model without systematic FIT failure increased 0.1 percentage point from the first to the second screening round (from 3.4% to 3.5%), while in real-life the false-positivity rate decreased 0.6 percentage point (from 3.5% to 2.9%). The simulated rate for the second round was outside the confidence interval of the observed rate. In the model with systematic FIT failure, the observed false positive rate was fitted best when assuming 1.3% of the participants had systematic false-positivity, which corresponds to 30% of the individuals with a false-positive result in the first round.

The detection rate for non-advanced adenomas simulated by the model without systematic FIT failure did not change between the first and second round (1.7% both rounds), while the observed detection rate declined from 1.7% to 1.2%. Here also, the second round simulated rate was outside the confidence interval of the observed rate. When assuming that 73% of non-advanced adenomas systematically tested false-negative, the model with systematic FIT failure fitted the observed second round detection rate best.

For advanced adenomas, the second round detection rate simulated without systematic FIT failure was inside the confidence interval of the observed rate, although again, the observed rate showed a larger decrease between the rounds (from 2.8% to 1.9 (simulated) versus 1.6% (observed)). In the model with systematic FIT failure, the observed detection rate was fitted best when assuming 26% of advanced adenomas systematically tested false-negative.

For colorectal cancer, the simulated detection rate declined 0.21 percentage point from the first to the second round (from 0.49% to 0.28%), while the observed decline was larger: 0.31 percentage point (from 0.49% to 0.18%). However, the second round cancer detection rate was inside the wide confidence interval of the observed rate.

Table 1: **Observed and simulated positivity rates, detection rates and the positive predictive value of the first and second round: data of the 1-sample immunochemical fecal occult blood test (FIT) with a cut-off of 10 µg Hb/g feces in all groups of the CORERO trial, and model without and with systematic FIT failure.**

	1st round			2nd round		
	Observed (95% CI) N=4523	Simulated without systematic FIT failure	Simulated with systematic FIT failure	Observed (95% CI) N=3427	Simulated without systematic FIT failure	Simulated with systematic FIT failure
*Positivity rate (%)**	8.4 (7.6-9.2)	8.3	8.3	5.8 (5.1-6.6)	7.3	5.8
*False-positive rate (%)**	3.5 (2.9-4.0)	3.4	3.4	2.9 (2.3-3.4)	3.5	2.7
*Detection rate of non-advanced adenomas (%)**	1.7 (1.3-2.0)	1.7	1.7	1.2 (0.9-1.6)	1.7	1.3
*Detection rate of advanced adenomas (%)**	2.8 (2.3-3.3)	2.7	2.7	1.6 (1.2-2.0)	1.9	1.6
*Detection rate of colorectal cancer (%)**	0.49 (0.28-0.69)	0.49	0.50	0.18 (0.04-0.32)	0.28	0.27
*Detection rate of advanced neoplasia (%)**	3.3 (2.8-3.8)	3.2	3.2	1.8 (1.3-2.2)	2.1	1.8
*Positive predictive value (%)**	38.9 (34.0-43.9)	38.5	38.6	30 (23.6-36.4)	29.4	31.7

* Number of cases divided by the total number of screenees.

Table 2: **Positivity rates, detection rates and the positive predictive value of the third round with 1-sample FIT and the first round with 2-sample FIT with a cut-off 10 µg Hb/g feces, as observed in group IV of the CORERO trial and in the model without and with systematic FIT failure.**

	1st round 2-sample FIT			3rd round 1-sample FIT		
	Observed (95% CI) N=1876	Simulated without systematic FIT failure	Simulated with systematic FIT failure	Observed (95% CI) N=2907	Simulated without systematic FIT failure	Simulated with systematic FIT failure
*Positivity rate (%)**	12.7 (11.2-14.3)	15.1	13.8	4.6 (3.9-5.4)	6.6	5.0
*False-positive rate (%)**	6.1 (5.0-7.2)	6.4	5.8	2.3 (1.8-2.9)	3.4	2.6
*Detection rate of non-advanced adenomas (%)**	2.6 (1.8-3.3)	3.3	2.9	1.3 (0.9-1.7)	1.6	1.1
*Detection rate of advanced adenomas (%)**	3.4 (2.6-4.2)	4.6	4.3	0.9 (0.6-1.3)	1.4	1.1
*Detection rate of colorectal cancer (%)**	0.69 (0.32-1.07)	0.75	0.75	0.10 (0.0-0.3)	0.21	0.22

* Number of cases divided by the total number of screenees.

Third round and two-sample group

The observed positivity and detection rates of the third round of 1-sample screening and the first round of 2-sample screening are shown in Table 2. Without systematic FIT failure, the simulated positivity and detection rates were all higher than observed and most fell outside the confidence intervals of the observed rates in the CORERO-trial. With systematic FIT failure, all the simulated positivity and detection rates decreased and fell within the confidence interval of the observed rates, although most remained higher than the observed rates.

Effectiveness

The introduction of systematic false-positive test results decreased the number of screenees that needed to undergo a colonoscopy to prevent one death: the NNScope was reduced by 7.4% (from 41.1 to 38.0), while maintaining almost 99% of life-years gained (Table 3). It reduced the average number of diagnostic and total colonoscopies done per person that started the screening program by 24% and 19%. With the introduction of systematic false-negative test results, incidence reduction, mortality reduction and life years gained (LYG) from screening declined with 9.4%, 4.5%, and 3.8% respectively compared to no systematic false-negativity, while also the NNScope of the program decreased. Together both elements of systematic FIT failure resulted in a decline of 13.0% in incidence reduction (from 40.9% to 35.6%), a decline of 6.4% in mortality reduction (from 59.0% to 55.2%), and a decline of 5.2% in LYG (from 245 to 232 per 1000 participants). If the LYG were 3% discounted, the decline was 4.8% (from 115 LYG to 109 LYG). The NNscope and the average number of diagnostic and total colonoscopies per person also declined. When we assumed that all first round false-negative lesions were systematically false-negative, the incidence reduction now declined with 34.8% from 40.9% to 26.7%, the mortality reduction declined with 16.4% from 59.0% to 49.3% and the LYG declined with 14.0% from 245 to 211 per 1000 participants.

Table 3: **Simulated changes in life long outcomes of the screening program (55-75 years, 2-year interval, immunochemical fecal occult blood test with cut-off 10 µg Hb/g feces) in participants, four FIT failure model versions, no discounting.**

	Model without systematic FIT failure	Model with systematic false-negative test results	Model with systematic false-positive test results	Model with systematic FIT failure (both types)
Incidence reduction	40.9%	37.1%	39.8%	35.6%
Mortality reduction	59.0%	56.4%	58.1%	55.2%
Average # colonoscopies per participant (diagnostic + surveillance)	0.86	0.79	0.78	0.7
Average # of diagnostic colonoscopies per participant	0.49	0.39	0.44	0.37
*NNscope to prevent one death**	41.1	39.7	38	36.4
Life years gained[†]	245	235	242	232
QALYs gained[‡]	257	243	253	238

* NNScope = number needed to scope.
[†] Per 1000 persons starting the screening program.
[‡]QALYs = Quality adjusted life years

DISCUSSION

Our current analysis shows that the lower detection rate of adenomas observed in the second round of FIT screening is not only caused by depletion of the adenoma prevalence pool. We needed to assume that 73% of non-advanced adenomas and 26% of advanced adenomas are systematically missed by FIT to simulate the observed decrease in detection rates between the first and the second round FIT screening. We also needed to assume that 30% of false-positives are systematically false-positive to simulate the observed decrease in positivity rate. Compared to the projections without systematic FIT failure, the projections with systematic FIT failure resulted in a 5.2% decrease of life-years gained by biennial FIT screening from age 55 to 75 years, while the incidence and mortality reduction decreased 13.0% and 6.4% respectively.

Systematic false-negative test results are a very plausible and even expected phenomenon. Non-bleeding adenomas would be the most obvious explanation for this phenomenon. Other explanations include a longer time for hemoglobin decay caused by a proximal location of adenomas or longer colonic transit in a person, and different sampling techniques. It is expected that systematic false negative FIT results are associated with decreased effectiveness of a FIT screening program, since adenomas that are systematically missed in consecutive screening rounds decrease the program sensitivity of a multi-round screening program. This impairs screening effectiveness in terms of incidence and mortality reduction.[18] Systematic false-positive test results are

also not unexpected. Obviously, there are individuals with a constitution or (chronic) condition causing fecal bleeding. If those individuals are selected out at earlier rounds, it explains why the number of false-positive test results in subsequent rounds decrease. Accounting for this latter mechanism had a positive influence by reducing the NNScope and the average number of colonoscopies per participant. The NNscope to prevent one death was also reduced when accounting for systematic false-negative test results. This may seem surprising at first, but is fully plausible. As shown by the calibrated percentages, systematic false-negative test results especially play a role in non-advanced adenoma, where they reduced the number of colonoscopies relatively more than that they reduced the effectiveness of adenoma removal.

In this analysis, we focused on systematic false-negativity of FIT for the detection of adenomas. Based on an earlier analysis of the gFOBT hemoccult II trials, we already identified a lower sensitivity for carcinomas long before clinical diagnosis compared to shortly before clinical diagnosis, which we since then incorporated in our model.[17] Due to the low number of carcinomas in the CORERO-trial it was not possible to reliably validate these earlier assumptions regarding cancers.

A strength of this study is that we validated the estimated systematic FIT failure to the third round of screening as well as in a different study population using 2-sample FIT. Here also the model with systematic FIT failure better matched real-life observations than the model without, both for non-advanced adenomas and advanced adenomas. This further strengthens the case for occurrence of systematic false-negative FIT results.

Two limitations are noteworthy. Firstly, in this study we estimated systematic FIT failure for a cut-off level of 10 µg Hb/g feces. Based on the available data, we could not estimate it for higher cut-offs. However, if systematic false-negative FIT results are primarily caused by non-bleeding adenomas, one could argue that the same number of adenomas will not bleed when testing with a higher FIT cut-off. Then the differences in detection rates between cut-offs will fully be explained by the sensitivity for adenomas that do bleed. Secondly, we assumed that bleeding and non-bleeding adenomas have the same probability to progress to cancer. If non-bleeding adenomas have a lower probability to progress than bleeding adenomas, this will attenuate the impact on the effectiveness of FIT screening.

To our knowledge there is no other colorectal cancer screening model that addressed the issue of systematic false-negative FIT results to compare our estimates with. Regarding observed data of consecutive screening rounds, an earlier FIT study in Italy reported an even larger difference in (advanced) adenoma yield between the first and the second round after 2 years than observed from the CORERO-trial data we used.[21] On the other hand, a Dutch FIT study in Amsterdam reported a less pronounced difference between the rounds.[19] Given these data and the confidence intervals of

the rates in the CORERO-trial, there is still uncertainty around the exact amount of systematic false-negative FIT results. It is therefore interesting that under the extreme assumption that all adenomas missed at first round will be missed at subsequent rounds, the mortality reduction reduced by 16.4% (from 59.0% to 49.3%) instead of 6.4% (from 59.0% to 55.2%). There is also some uncertainty about the exact mechanism. We assumed that adenomas that have never bled so far, can only start to bleed when growing from a non-advanced to an advanced stage. In real life this will happen on a more continuous basis, for example when growing from 6 to 8 mm. We also assumed adenomas have either no sensitivity or full sensitivity, while in real life presumably, some adenomas will bleed more often or may shed more blood when bleeding than others, leading to variable sensitivity levels.

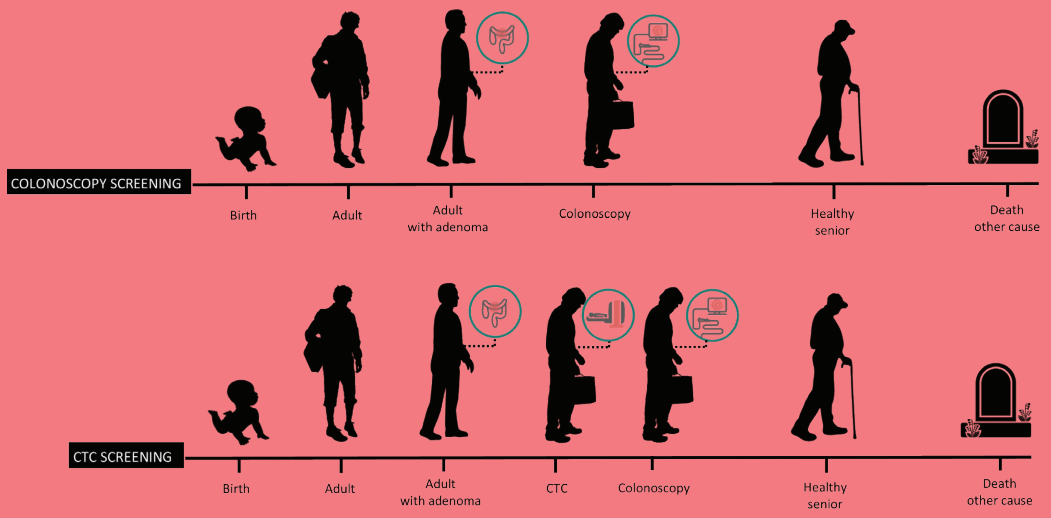
In conclusion, this study provides convincing evidence that a proportion of adenomas is systematically missed by repeated FIT screening, presumably due to non-bleeding adenomas. This phenomenon lowers the impact of FIT screening on mortality reduction with an estimated 6.4%.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin.* 2011;61(2):69-90.
2. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. *Cancer.* 1986;58(2):397-403.
3. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet.* 1996;348(9040):1467-1471.
4. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst.* 1999;91(5):434-437.
5. Allison JE, Sakoda LC, Levin TR, et al. Screening for colorectal neoplasms with new fecal occult blood tests: update on performance characteristics. *J Natl Cancer Inst.* 2007;99(19):1462-1470.
6. Guittet L, Bouvier V, Mariotte N, et al. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. *Gut.* 2007;56(2):210-214.
7. 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.
8. 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):1103-1110.
9. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. *Am J Gastroenterol.* 2010;105(9):2017-2025.
10. 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.
11. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2008;149(9):638-658.
12. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut.* 2015;64(10):1637-1649.
13. Gupta S, Sussman DA, Doubeni CA, et al. Challenges and possible solutions to colorectal cancer screening for the underserved. *J Natl Cancer Inst.* 2014;106(4):dju032.
14. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. *Epidemiol Rev.* 2011;33(1):88-100.
15. Lang CA, Ransohoff DF. On the sensitivity of fecal occult blood test screening for colorectal cancer. *J Natl Cancer Inst.* 1997;89(19):1392-1393.
16. Ransohoff DF, Lang CA, Young GP. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology.* 1997;113(4):1423-1424.
17. Lansdorp-Vogelaar I, van Ballegooijen M, Boer R, Zauber A, Habbema JD. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. *Cancer.* 2009;115(11):2410-2419.

18. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. Impact of systematic false-negative test results on the performance of faecal occult blood screening. *Eur J Cancer.* 2001;37(7):912-917.
19. Denters MJ, Deutekom M, Bossuyt PM, Stroobants AK, Fockens P, Dekker E. Lower risk of advanced neoplasia among patients with a previous negative result from a fecal test for colorectal cancer. *Gastroenterology.* 2012;142(3):497-504.
20. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. *Gut.* 2013;62(3):409-415.
21. Zorzi M, Barca A, Falcini F, et al. Screening for colorectal cancer in Italy: 2005 survey. *Epidemiol Prev.* 2007;31(2-3 Suppl 2):49-60.
22. van Roon AH, Wilschut JA, Hol L, et al. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. *Clin Gastroenterol Hepatol.* 2011;9(4):333-339.
23. 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):1257-1264.
24. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res.* 1999;32(1):13-33.
25. Vogelaar IvB, M.; Zauber, A.G. *Modeler Profiler of the MISCAN-Colon Microsimualtion Model For Colorectal Cancer. .* Department of Public Health, Eramus Medical Center.
26. Morson B. President's address. The polyp-cancer sequence in the large bowel. *Proc R Soc Med.* 1974;67(6 Pt 1):451-457.
27. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. *Cancer.* 1975;36(6):2251-2270.
28. Comprehensive Cancer Centre. http://cijfersoverkanker.nl/. Accessed 14 Mar 2012.
29. Arminski TC, McLean DW. Incidence and Distribution of Adenomatous Polyps of the Colon and Rectum Based on 1,000 Autopsy Examinations. *Dis Colon Rectum.* 1964;7:249-261.
30. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. *Cancer.* 1988;61(7):1472-1476.
31. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. *Ann Surg.* 1963;157:223-226.
32. Clark JC, Collan Y, Eide TJ, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. *Int J Cancer.* 1985;36(2):179-186.
33. Jass JR, Young PJ, Robinson EM. Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. *Gut.* 1992;33(11):1508-1514.
34. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. *Scand J Gastroenterol.* 1989;24(7):799-806.
35. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. *Cancer.* 1979;43(5):1847-1857.
36. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. *Cancer.* 1982;49(4):819-825.
37. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut.* 1982;23(10):835-842.
38. 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.

39. Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet.* 2010;375(9726):1624-1633.
40. van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. *Am J Gastroenterol.* 2006;101(2):343-350.


COLONOSCOPY SCREENING
Birth
Adult
Adult with adenoma
Colonoscopy
Healthy senior
Death other cause
CTC SCREENING
Birth
Adult
Adult with adenoma
CTC
Colonoscopy
Healthy senior
Death other cause

Chapter 5

Cost-effectiveness of colonoscopy versus CT-colonography screening for colorectal cancer with observed participation rates and costs.

Miriam P. van der Meulen, I. Lansdorp-Vogelaar, S.L. Goede, E.J. Kuipers , E. Dekker, J. Stoker, M. van Ballegooijen

Radiology. 2018;287(3):901-11.

ABSTRACT

Purpose: The aim of this study was to determine the comparative cost-effectiveness of CTC versus colonoscopy screening using data on unit costs and participation rates from a randomized controlled screening trial in a dedicated screening setting.

Materials and Methods: Observed participation rates and screening costs in a randomized controlled screening trial were used in a microsimulation model to estimate costs and quality adjusted life years (QALYs) gained of colonoscopy and CTC screening. For both tests we determined optimal age range and interval combinations assuming a 100% participation rate. Assuming observed participation for these combinations, we compared the cost-effectiveness of both tests. Extracolonic findings were not included as long-term follow-up data are lacking.

Results: The participation rates for colonoscopy and CTC were 21.5% (1276/5924) and 33.6% (982/2920) respectively. Colonoscopy was more cost-effective in the screening strategies with one or two lifetime screens, whereas CTC was more cost-effective in strategies with more lifetime screens. CTC was the preferred test for willingness-to-pay-thresholds of €3,200 per QALY gained and higher, which is lower than the Dutch willingness-to-pay threshold of €20,000. With equal participation, colonoscopy was the preferred test independent of willingness to pay thresholds. The findings were robust for most of the sensitivity analyses, except with regard to relative screening costs and subsequent participation.

Conclusion: Because of the higher participation rates, CTC screening for colorectal cancer is more cost-effective than colonoscopy screening. The implementation of CTC screening requires prior satisfactory resolution to the question as to how best to deal with extracolonic findings.

INTRODUCTION

Colorectal cancer (CRC) is the second-most common cause of cancer mortality in the western world.(1) Early detection and treatment of CRC can reduce CRC incidence and mortality. Several randomized controlled trials have demonstrated a mortality reduction with the use of guiac fecal occult blood test (gFOBT)(2-5) and sigmoidoscopy.(6)

Colonoscopy is considered the reference standard and allows for direct removal of colonic lesions. There is only indirect evidence of the preventive effect, however, and its use as a primary screening test has drawbacks, including its perceived burden, low rates of patient compliance with screening recommendations, risk of complications, costs, and need for endoscopy capacity.(7) This burden can be reduced by selecting individuals with increases risk to undergo colonoscopy, for instance by incorporating a less invasive primary test and offer colonoscopy only to test positives. CT-colonography (CTC) is an example of such a test, with the benefit that it is a less invasive screening method, needs less extensive bowel preparation,(8) and that it has a lower risk of complications.(9)

In several studies the accuracy and one-round yield of advanced neoplasia was slightly higher with colonoscopy than with CTC.(10-12) In addition, most of the comparative cost-effectiveness studies showed that colonoscopy screening strategies were more cost-effective than CTC strategies when assuming full attendance.(13-15) However, the outcomes of these analyses were highly sensitive to the participation rates and costs of the two screening tests.

With a lower participation rate, screening is applied in fewer individuals, decreasing both the effectiveness and costs. Therefore, differences in participation between tests will influence their relative cost-effectiveness. In 2012, a large randomized controlled trial in the Netherlands was the first to compare the participation, yield and unit costs of CTC and colonoscopy in a dedicated screening setting.(12) Participation was indeed higher for CTC (33.6%, 982/2920) than for colonoscopy (21.5%, 1276/5924),(12) while unit costs for CTC were 25% lower than those of a negative colonoscopy and 53% lower than those of a positive colonoscopy (because of additional costs for colonoscopy with polypectomy).(16, 17) These estimates allowed for a more representative cost-effectiveness analysis of CTC and colonoscopy screening. Therefore, the aim of this study was to determine the comparative cost-effectiveness of CTC versus colonoscopy screening based on the data from a randomized controlled screening trial in a dedicated screening setting.

MATERIALS AND METHODS

The COCOS-trial

Costs and participation data were obtained from the COCOS-trial, a randomized controlled trial of CTC versus colonoscopy screening.(12, 16, 17) In this trial, screening-naïve members of the general population aged 50-75 years and living in the regions of Amsterdam and Rotterdam were identified via the regional municipal administration registries and randomly allocated (2:1) to invitation to primary screening by colonoscopy or CTC. Participation to the first screening round was 21.5% (1276/5924) for colonoscopy versus 33.6% (982/2920) for CTC.(12) A positive CTC was estimated to cost €158; a negative CTC €149; a positive colonoscopy (including polypectomy and pathology costs) €329; and a negative colonoscopy €192.(16, 17) Details on bowel preparation, protocols and yield are given in the extended methods section (Appendix 1).

Microsimulation model

We used the MISCAN-Colon model which is a previously described microsimulation model for CRC ((18, 19) and the Model Appendix. A microsimulation model can incorporate trial results to estimate the lifetime effects of findings of the randomized trial. Additionally, the model can estimate the cost-effectiveness of various screening strategies instead of only the screening strategy used in the randomized trial and thereby study which screening strategies are optimal.

This model simulates the life histories of a large population of individuals from birth to death. As a simulated person ages, one or more adenomas may develop and may progress in size from small (≤5mm) to medium (6-9mm) to large (≥10mm). Some adenomas can develop into preclinical cancer, which may progress through stages I to IV. At any time, the process may be interrupted by death from another cause.

Screening may prevent the development of cancers as it allows detection and removal of adenomas and it may detect CRC in an earlier stage with a more favorable survival. Thus, CRC incidence and mortality could be reduced. The life years gained by screening are calculated as the difference in model-predicted life years lived in the population with and without CRC screening. The natural history is further described in the Model Appendix.

Study population

In this study, we modelled the age distribution of the Dutch population aged 25 to 85 years in 2015 (20) and followed all individuals until death. Life-expectancy was based on sex-specific life-tables from 2011 obtained from Statistics Netherlands.(20)

Screening strategies

Screening started in 2015. We simulated four different strategies for referral of CTC positive individuals to diagnostic colonoscopy, see Figure 1.

Individuals were offered different screening schedules varying by:

- Age to start screening: 40, 45, 50, 55, 60 and 65 years
- Age to stop screening: 70, 75, 80 and 85 years
- Screening interval: 3, 5, 10, 15 and 20 years

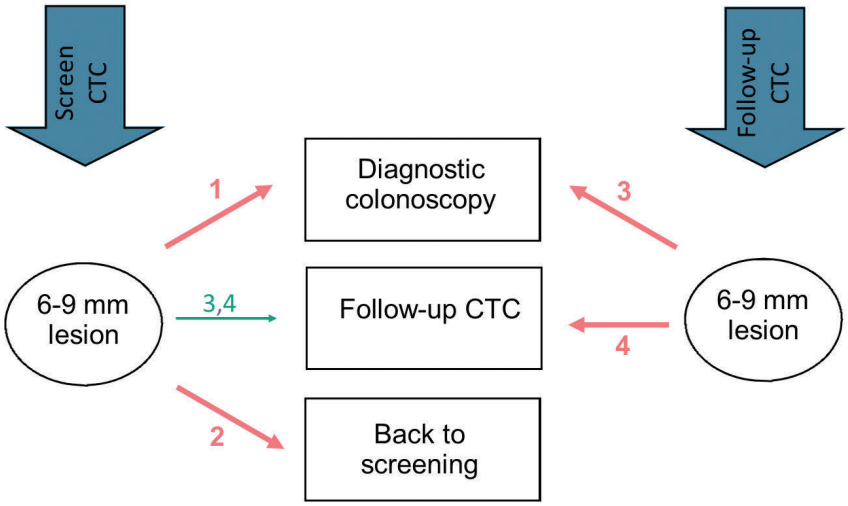
Together, the various screening ages and intervals resulted in 86 screening strategies for colonoscopy and each follow-up CTC strategy, combined to a total of 480 strategies.

If adenomas were detected during primary or diagnostic colonoscopy, they were removed and the individual entered surveillance according to the Dutch guidelines.(21)

Test characteristics and size distribution of adenomas

We assumed the sensitivity of colonoscopy to be 75% for adenomas with a diameter of 1-5mm, 85% for adenomas 6-9mm, and 95% for adenomas ≥10mm and CRC.(22) The specificity of colonoscopy was assumed to be 90% to account for the presence of non-adenomatous polyps. We assumed the sensitivity of CTC to be 75.7% for adenomas sized 6-9mm, 85.9% for adenomas ≥ 10 mm and 95% for colorectal carcinoma. Specificity was assumed to be 91.4% with a cut-off of 6mm and 97.6% with a cut-off of 10mm.(10, 23) We validated the model with these test characteristics with observations in the COCOS-trial (see Appendix 2), and adjusted the size distribution of adenomas (but not the prevalence) in the model to fit the data best.

Figure 1: **The four different simulated follow-up strategies for CTC screening which differ when 6-9 mm adenomas were found.**



In all four strategies, individuals with lesions of ≥10mm at CTC were immediately referred for diagnostic colonoscopy, and individuals without lesions or with lesions 1-5mm at CTC returned to the screening program. The follow-up strategies differed in the management of medium sized lesions (6-9mm) seen at CTC. Individuals with medium-sized lesions 1) were directly referred for diagnostic colonoscopy (i.e. using a cut-off of 6mm); 2) returned to the screening program (i.e. corresponding with a 10mm cut-off); 3) and 4) were offered a follow-up CTC after 3 years, as was done in the COCOS-trial and referred to a diagnostic colonoscopy if they at follow-up CTC had 3) a medium or large lesion, or 4) a large lesion. In the 4th strategy, persons with medium-sized adenomas continued to receive follow-up CTC, either until a large lesion was detected or until a medium lesion was no longer detected.

Participation

We first assumed a 100% participation rate to identify the optimal screening strategies in terms of interval, age range and CTC cut-off. Subsequently, we simulated the optimal screening strategies with the observed participation rates to compare the two screening tests in terms of cost-effectiveness

We assumed 100% participation to determine the optimal strategies, because assuming observed participation could result in optimal screening strategies for the population which are at too short intervals. These short intervals in the population will result in average screening intervals at the most optimal length for an individual. However, this would lead to overscreening in those who adhere with recommendations, which we feel is unethical and which in turn might also lead to lower compliance in practice than assumed in the analysis.

To compare the two screening tests in terms of cost-effectiveness, we then simulated the optimal screening strategies with the observed participation rates (21.5% (1276/5924) for colonoscopy versus 33.6% (982/2920) for CTC). The observed

age-dependency and participation to diagnostic colonoscopy after positive CTC were modelled accordingly. Participation to surveillance colonoscopies was assumed to be 80%.(24) In a sensitivity analysis, we also modeled all possible screening strategies with observed participation.

Data on subsequent participation to colonoscopy or CTC screening are lacking. We assumed stable overall participation to subsequent rounds and assumed that 90% of the previous responders also attended the subsequent screening round, as found in a Dutch FIT screening trial.(25) To maintain stable overall participation, the remaining percentage was filled with previous non-attenders. In addition, we assumed that 10% of the individuals never attended screening (24) and that the risk of CRC in this group was higher than that in the general population (RR=1.15).(2)

Costs and utility losses

We included all costs from a third-party payer perspective. We assumed unit costs for CTC and colonoscopy as observed in the COCOS-trial. We further included screening and treatment costs and utility losses as presented in Table 1.

Extracolonic findings

Extracolonic findings were not included in the cost-effectiveness analysis as long-term follow-up data are lacking. We did, however, include a sensitivity analysis accounting for the currently available data (see Appendix Figure 3b). Two scenarios were simulated; in the first only extra costs and utility loss due to extracolonic findings were accounted for. In the second scenario we also assumed a benefit of the detection and further management of abdominal aortic aneurysms (AAA).

Outcomes

We used the model to predict costs and QALYs gained for all screening strategies compared to no screening. Costs and QALYs gained were discounted by 3% per year in accordance with international literature.(32) Alternatively, Dutch discounting (1.5% for QALYs and 4% for costs) was used in a sensitivity analysis.

Analysis

We first showed disaggregated outcomes behind the total costs and QALYs gained finally used in the CEAof screening every 10 years from age 50-70 for colonoscopy and CTC screening with a 6 mm cut-off.

Next, the total costs and QALYs gained of all 480 simulated strategies were compared. Per test, it was determined which strategies were efficient by ruling out strategies that were more costly and less effective than other strategies (simple

dominance) or combinations of other strategies (extended dominance). On a plot of QALY's gained versus costs, the line that connects the efficient strategies is called the efficient frontier. The incremental cost-effectiveness ratio (ICER) of each efficient strategy was calculated by comparing its costs and effects with those of the next less costly and less effective efficient strategy. An ICER of less than €20,000 per QALY gained was assumed to be cost-effective.(33)

Sensitivity analysis

In addition to the above-mentioned sensitivity analysis, we performed one-way sensitivity analyses on several other parameters, summarized in Table 1.

RESULTS

Screening every 10 years

With a 100% participation rate, 10-yearly screening from age 50-70 resulted in a higher mortality reduction and more QALYs gained with colonoscopy than with CTC (106 vs 81 QALY's gained, 24% lower) (Table 2). With colonoscopy screening, the screening and surveillance costs were higher than with CTC, but the CRC treatment savings were also higher. This resulted in lower total costs of the colonoscopy screening program. Therefore, with this screening strategy, colonoscopy dominated CTC screening.

With observed participation, CTC screening resulted in a higher mortality reduction, more QALY's gained (29 versus 22 QALYs gained, 34% higher), but still higher total costs. The number of lifetime colonoscopies and complications and the number of people needed to scope to detect one advanced neoplasia were lower with CTC screening assuming both 100% and observed participation.

Optimization of interval, age range and CTC cut-off

When all strategies were considered with a 100% participation rate (varying screening age range and interval), only colonoscopy strategies appeared on the efficient frontier (see Figure 2). The optimal colonoscopy and CTC strategies are presented in Table 3. The most effective colonoscopy strategy with an ICER below the Dutch threshold of €20,000 per QALY gained was colonoscopy every 5 years from age 50 to 70 (strategy 4 in Table 3).

Of the CTC strategies, those with a 6mm cut-off resulted in more QALY's gained and lower total costs than other follow-up strategies. The most effective CTC strategy with an ICER below the Dutch threshold of €20,000 per QALY gained was CTC with a 6 mm cut-off every 3 years from age 45 to 80 (strategy 8 in Table 3).

Table 1: **Summary of model assumptions of the base case and sensitivity analyses**

Variable	Base case analysis		Sensitivity analyses	Source
	Colonoscopy	CTC		
Test characteristics				
sensitivity				
1-5 mm adenomas	75%	0%	0.5* (1-sensitivity) systematically missed in CTC alone and in both CTC and colonoscopy	Colonoscopy: systematic review(22)
6-9 mm adenomas	85%	75.7%/0%		
≥10 mm adenomas	95%	85.90%	Proximal 50% less sensitivity in colonoscopy alone and in both CTC and colonoscopy	CTC: meta-analysis(10)
carcinomas	95%	95%	CTC 10 mm cut-off 7.6% sensitivity 6-9 mm adenomas	
specificity	90%*	91.4%/97.6%		
Participation	**Depending on age**		**Not depending on age**	**COCOS-trial(12)**
Screening test	100%/21.5%	100%/33.6%		
Subsequent participation	Plausible: previous-participants: 90%		Fixed: previous participants: 100%, Random: previous participants: 21.5%/33.6%	FIT-pilot(25)
Diagnostic test	na	100%/100%		
Surveillance test	100%/80%	100%/80%		US RCT(24)
Quality of life loss†				
Test	0.5 for 1.5 day	0.5 for 1.5 day	2 days	
Complication	14 days		-	Expert opinion
CRC from diagnosis onward	Initial treatment stage 1 till IV: 0.12; 0.18; 0.24; 0.70		50%/200%	
	Continuous care stage 1 till IV: 0.12; 0.18; 0.24; 0.70			Survey(26)
	Terminal care death by CRC: 0.70			
	Terminal care death by other cause: 0.12; 0.18; 0.24; 0.70			
Fatal complications after colonoscopy	**3.29^10^{5} in positive colonoscopies**		**50%/200%**	
Screening costs†‡				
Screening costs (positive/negative)	€329‡/€192	€158‡/€149	50%/200%	COCOS-trial(16, 17)
Costs nonresponder	€ 6	€ 6		
Diagnostic costs inside screening program	na	€329/€192	50%/200%	

Table 1: **Summary of model assumptions of the base case and sensitivity analyses**

Variable	Base case analysis		Sensitivity analyses	Source
	Colonoscopy	**CTC**		
Costs complications after colonoscopy	€ 5,100		50%/200%	Estimated from US cost-analysis study(27)
Treatment costs[†‡]	*Initial Continuous Terminal care Terminal care treatment care death CTC death other c cause*			**Initial treatment and continuous care: Dutch DTC rates (28, 29)**
stage I	€13,773	€375 €19,282 €4,848	50%/200%	Terminal care death CRC: Dutch last year of life cost-analysis (30)
stage II	€18,180	€375 €19,282 €4,407		
stage III	€20,935	€375 €20,384 €5,729		Terminal care death other cause: Estimated from US cost-analysis (31)
stage IV	€27,546	€375 €27,546 €15,426		
Size distribution of adenomas[$]	Based on colonoscopy studies		Based on autopsy studies	
Discounting	3%		No discounting/ 1.5% discounting on QALY's and 4% on costs	International guidelines(32)

* To account for the presence of hyperplastic polyps and associated risks and costs of removal
[†]Further described in Appendix 1
[‡]All costs were inflation adjusted to 2012 euro's using the Dutch Consumer Price Index; Costs for a positive colonoscopy included polypectomy and pathology costs; A positive CTC has slightly higher costs than a negative CTC because of the costs for consultation after the positive test result.
[$]We validated the model with observations in the COCOS-trial, which is shown in Appendix 2, and adjusted the size distribution of detected adenomas in the model, while we previously based the size distribution on autopsy studies.

Comparing CTC and colonoscopy

When the optimal colonoscopy and CTC strategies were simulated with observed participation, both costs and QALYs gained decreased compared to full participation (Figure 3, Table 3). The costs and QALYs gained assuming observed participation rates decreased more for colonoscopy screening than for CTC screening. Still, colonoscopy screening with one or two lifetime screens was less costly than and just as effective as the same CTC strategies. However, with more lifetime screens, CTC screening dominated the colonoscopy strategies. The first CTC strategy on the efficient frontier had a screening age of 55-70 years, an interval of 5 years and an ICER of €3162 per QALY gained. The most effective CTC strategy with an ICER below the €20,000 threshold was CTC triennially from age 45 to 80 years (ICER €14,709 per QALY gained).

Sensitivity analyses

In concordance with the base case analyses, most sensitivity analyses showed that colonoscopy with one or two lifetime screens dominated CTC screening, while with more lifetime screens, CTC dominated the colonoscopy strategies (Appendix 3a). Simulating all screening strategies with observed participation did not change which screening strategies were efficient and produced the same pattern. Including extracolonic findings made CTC less cost-effective in the version without and more cost-effective in the version with benefits of detected AAA, while CTC still dominated colonoscopy screening in both cases, showing an ICER of €3478 and €2458 per QALY gained, respectively (Appendix 3b). If CTC costs were doubled or colonoscopy costs halved, colonoscopy dominated CTC and vice versa (Appendix 3c, Panel I-N). Also, when subsequent participation was random, colonoscopy dominated CTC screening (Appendix 3c, Panel V).

Table 2: **Modeling outcomes of a screening program with colonoscopy or CT-colonography with 6 mm cut-off (strategy 1 in figure 1) every 10 years from age 50 to 70 with a 100% participation rate and using observed participation.**

	Colonoscopy		CT-colonography 6 mm cut-off	
	100% attendance	observed attendance	100% attendance	observed attendance
Incidence reduction*	50.0%	10.2%	33.9%	11.8%
Mortality reduction*	60.4%	12.3%	46.4%	16.3%
Deaths prevented*	25	5	20	7
QALY's gained*†	106	22	81	29
Life years gained*†	100	20	79	29
Life years in therapy*†	-120	-24	-29	-8
# persons ever screened		22.4%		42.6%
# persons with any detected adenoma	35.0%	7.6%	18.7%	7.4%
# persons with detected advanced adenoma	6.1%	1.3%	7.1%	2.9%
# persons with detected colorectal cancer	0.8%	0.2%	1.2%	0.5%
Lifetime complications*	7.3	1.5	4.3	1.4
Lifetime colonoscopies	2.9	0.6	0.7	1.0
Number Needed to Scope	27.6	27.7	2.6	2.5
Screening costs (€)*†‡	292070	68212	203961	82999
Diagnostic costs (€)*†‡	0	0	47123	17852
Surveillance costs (€)*†‡	124260	22264	59461	15774
Complications costs (€)*†‡	22952	4399	13218	4219
Treatment costs (€)*†‡	-361342	-72540	-236132	-81998
Total costs (€)*†‡	77941	22335	87631	38846

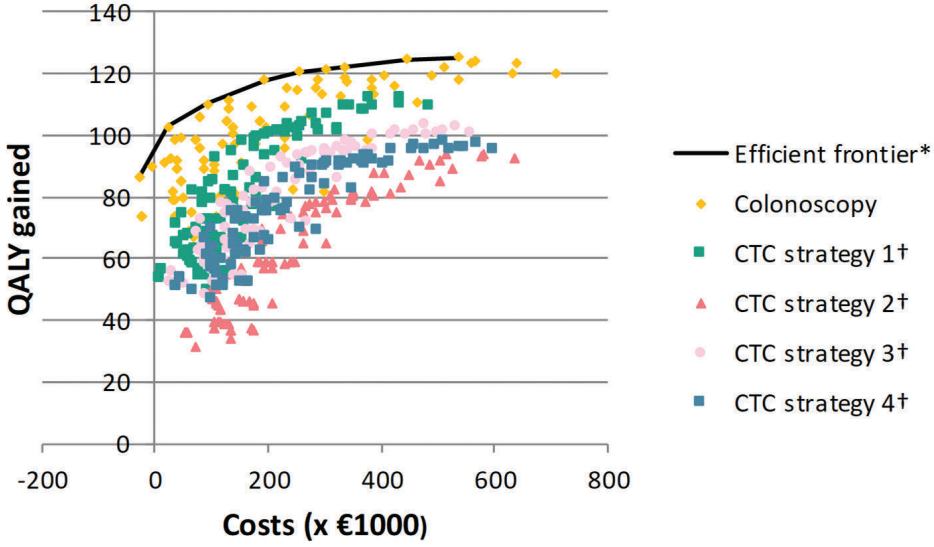
* per 1000 invitees
† 3% discounted
‡ Adjusted to 2012 euros

Table 3: **Modeled costs and QALY's gained of the efficient screening strategies a) for colonoscopy and CTC separate, assuming a 100% participation rate, and b) comparing colonoscopy and CTC, assuming observed participation, compared to no screening with 3% discounting.**

	Screentest	Start age	Stop age	Interval	# Screens	QALY gained*ǂ	Costs*ǂ (€)	Costs/QALY gainedǂ	ICER
100% ATTENDANCE									
Colonoscopy									
1	Colonoscopy	60	na	na	1	86	-27,831	-323	-323
2	Colonoscopy	55	65	10	2	103	23,344	227	3,078
3	Colonoscopy	45	65	10	3	110	92,110	837	9,520
4	Colonoscopy	50	70	5	5	118	193,569	1,642	12,912
5	Colonoscopy	45	70	5	6	121	253,439	2,100	21,766
6	Colonoscopy	45	70	3	9	125	443,456	3,558	47,857
7	Colonoscopy	40	70	3	11	125	537,414	4,291	154,271
CTC									
1	CTC 6mm cut-off	65	na	na	1	53	9,760	183	183
2	CTC 6mm cut-off	60	na	na	1	56	14,540	258	1,573
3	CTC 6mm cut-off	60	70	10	2	71	39,881	560	1,718
4	CTC 6mm cut-off	60	70	5	3	82	66,668	817	2,558
5	CTC 6mm cut-off	55	70	5	4	92	108,471	1,174	3,897
6	CTC 6mm cut-off	50	70	5	5	98	154,286	1,578	8,471
7	CTC 6mm cut-off	50	75	3	9	107	278,983	2,616	14,049
8	CTC 6mm cut-off	45	80	3	12	112	377,502	3,372	18,524
OBSERVED ATTENDANCE									
1	Colonoscopy	60	na	na	1	17	-2,544	-146	-146
2	Colonoscopy	55	65	10	2	24	9,208	390	1,883
3	CTC 6mm cut-off	55	70	5	4	36	47,667	1,332	3,162
4	CTC 6mm cut-off	50	75	3	9	47	123,725	2,644	6,899
5	CTC 6mm cut-off	45	80	3	12	50	165,129	3,328	14,709

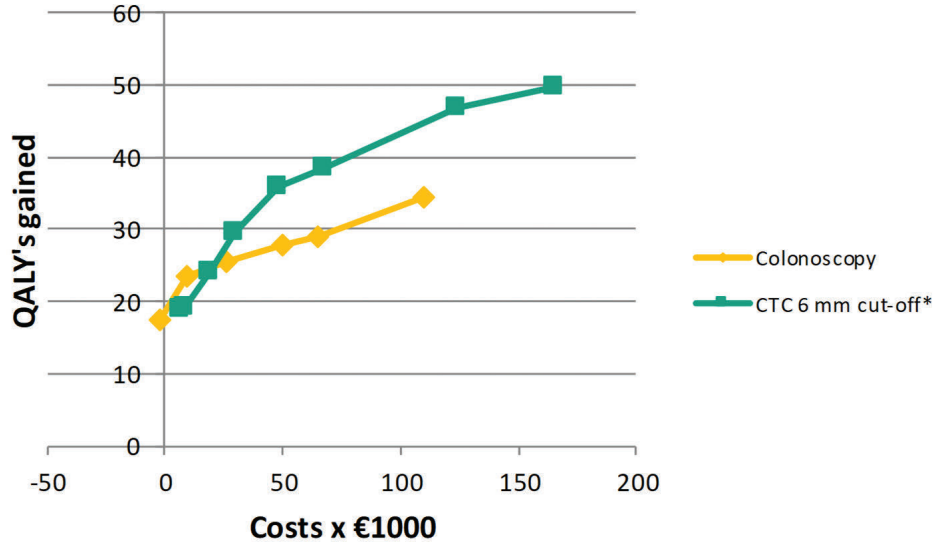
*per 1000 participants
ǂcompared to no screening

Figure 2: **Modeled costs and QALY's gained per 1000 participants for four CTC screening strategies and colonoscopy screening, with different starting and stopping ages and screening intervals, 3% discounted, from a participant's perspective.**



* Strategies on this frontier are presented in table 3a
† CTC strategy number correspond with numbers presented in Figure 1

Figure 3: **Modeled costs and QALY's gained per 1000 invitees of CTC strategies 6 mm cut-off and colonoscopy with different starting and stopping age and screening interval, 3% discounted, from a population's perspective.**



* CTC strategy 1 in figure 1

DISCUSSION

This study confirms that for people willing to participate in colorectal cancer screening, colonoscopy is more cost-effective than CTC screening. However, from a population's perspective, with participation as observed in the COCOS-trial, colonoscopy screening is less cost-effective than CTC screening if the latter offers more than two lifetime screens. Since the ICER of the least intensive CTC strategy on the cost-efficiency frontier was well below the Dutch threshold of €20,000 per QALY gained, CTC screening is preferred over colonoscopy screening as a one-test based screening program on a national level.

With observed participation and one lifetime screen, the lower sensitivity of CTC screening was compensated for by the higher participation, resulting in the same number of QALYs gained of CTC and colonoscopy. However, the costs of the diagnostic colonoscopy added to the CTC screening costs made CTC more expensive. The dominance of CTC with more lifetime screens can be explained by the yield of the additional screening rounds. After each screening round, the prevalence of (advanced) neoplasia in the screened population decreases. Since sensitivity for adenomas is higher for colonoscopy than for CTC, the residual number of adenomas is higher in CTC screenees than in colonoscopy screenees. Due to the higher participation rate, the number of CTC screenees is also higher than the number of colonoscopy screenees, further increasing the cost-effectiveness of intensifying screening in CTC compared to colonoscopy.

Previous studies assuming a 25% higher participation rate to CTC estimated that the costs of CTC should be no higher than 75%-95% of a colonoscopy to be more cost-efficient than colonoscopy.(34, 35) In our analysis, participation to CTC was 56% (33.6% versus 21.5%) higher compared to colonoscopy, while the costs of a negative and positive CTC were 75% (€ 149 of € 192) and 47% (€ 158 of € 329) of that of a negative and a positive colonoscopy, respectively. Thus, the relative participation rate of CTC was higher and relative costs of CTC lower than the studied threshold; therefore, our results were in line with these studies.

Since earlier costs-effectiveness analysis showed that their outcomes were highly sensitive to assumed participation rate and test costs,(13-15) an important strength of our study is that it is the first cost-effectiveness analysis that uses real-life data on unit costs and participation in a similar setting.

Four limitations of the study should be addressed. Lacking data on participation to subsequent screening rounds, we assumed participation for subsequent rounds was stable and assumed that 90% of previous responders also participated in a subsequent round.(25, 36-38) For the first aspect, although at this moment we have no concrete indication that participation will improve quickly, we acknowledge that with increased promotion, participation of colonoscopy and CTC might improve over the

years and participation of colonoscopy may than come close to that of CTC screening. The participation of colonoscopy should be quite close to that of CTC to dominate CTC screening (CTC should go down to 22.5%). For the latter aspect we explored two alternative scenarios in the sensitivity analysis, of which the outcome was that colonoscopy screening is preferred when assuming random subsequent participation, a scenario we find highly unlikely.

Second, we did not include two relevant aspects: exposure to ionizing radiation and extracolonic findings on CTC. Regarding radiation, a study estimated the ratio of colorectal cancers prevented to the number of radiation-related cancers induced at 24:1 to 35:1.(39) That study used a substantially higher ionizing radiation dose, however than the COCOS-trial and the studies we based our CTC sensitivity on. With respect to extracolonic findings, we conducted a sensitivity analysis with the sparse data on costs and abdominal aortic aneurysm screening, which showed a decrease in cost-effectiveness of CTC in the version without and an increase in cost-effectiveness in the version with benefit of detected AAA. As long-term follow-up of people with other screen-detected extracolonic lesions is lacking, the United States Preventive Services Task Force (USPSTF) concluded that the impact of extracolonic findings on the cost-effectiveness of CTC screening as yet cannot be estimated.(40) In theory, including extracolonic lesions could make CTC screening more cost-effective. On the other hand, detection of extracolonic lesions could be harmful, thereby reducing the (cost) effectiveness of CTC screening or even making CTC screening harmful in general. One could consequently argue that clinicians should ignore extracolonic findings, which could lead to other unacceptable ethical dilemmas. As long as this dilemma is unresolved, it is unknown whether CTC screening fulfils the WHO criteria that overall benefit should outweigh the harms.(41)

Third, we did not explicitly model distinct pathways for traditional and sessile serrated adenomas/polyps (SSA/P). The average time it takes for an adenoma to develop into CRC was calibrated to the randomized UK flexible sigmoidoscopy screening trial (42) and included both traditional adenomas and sessile serrated adenomas/polyps. Therefore, both adenoma types are included in the modeled mix of slow and rapid progressing lesions. An explicit separate pathway would be relevant if indeed CTC is less sensitive for SSA/P than colonoscopy (because they are often flat and therefore harder to detect with imaging(43)) and if the malignant potential of SSA/P towards cancer is different from the traditional pathway, which remains to be determined.

Although this cost-effectiveness analysis is primarily performed for the Dutch situation, we think that the results are relevant for other situations as well. Although absolute participation rates and costs may differ between countries, the relative differences between CTC and colonoscopy primarily determine the comparative

effectiveness. We suspect these relative differences in participation and costs differ less between countries. may be more similar. A recent Italian study on comparative participation rates of CTC and colonoscopy found a similar difference between the two tests as the COCOS-trial (colonoscopy 15% and versus CTC 25% and 28% depending on the level of cathartic preparation).(44) Furthermore, costs of CTC and colonoscopy consist for a large part of personnel costs. Because it is likely that there is a constant ratio of CTC/colonoscopy personnel costs, the ratio between unit costs for CTC and colonoscopy will not be influenced. Indeed, previous costs analyses found a similar ratio between CTC and colonoscopy Medicare reimbursement rates.(34) In contrast, a study by Pyenson et al. resulted in relatively lower costs for CTC, probably making CTC cost-effective also with equal participation rates.(45) Furthermore, the use of the Dutch willingness-to-pay threshold (€20,000 per QALY gained) does not imply that our findings cannot be generalized, because CTC showed an ICER of €3200 per QALY gained compared to colonoscopy and many countries have an even higher willingness-to-pay threshold than the Dutch threshold. Another issue concerning generalizability is that CTC is not being performed on a large scale yet. This could impact performance on community level. Indeed, a retrospective analysis in the English Bowel Cancer Screening Program and a Dutch study in which six physicians and three radiographers were trained both showed that performance of CTC was higher at centers and for radiologists with more experience.(46, 47) However, the Dutch study estimated the number of CTC examinations needed to achieve sufficient performance was only 164.(46) Currently, the target population for colorectal cancer screening in the Netherlands encompasses approximately 4.5 million individuals. Therefore, we conclude that experience to achieve an adequate performance of CTC in an organized screening program can be met in a short period of time.

The implications of our study depend on the way screening is introduced. Our study conclusions that CTC is preferred over colonoscopy is most applicable in a national screening program offering a single test modality (which is usually the case in so called organized screening programs). A comparison with alternative screening tests, however, should also be updated in future research, such as the fecal immunochemical test (FIT). Other options are also in place, mostly in opportunistic screening settings. An example of such an option is to offer participants a choice between screening tests. Since both colonoscopy and CTC are cost-effective in participants when compared to no screening, either test might be offered. Another option is to subsequently offer, different screening modalities to non-participants, starting with the most cost-effective screening test for participants. (48) Colonoscopy could then be offered initially, while non-participants could be offered CTC (or FIT). Three randomized controlled CRC screening trials that included a study arm offering a choice between screening tests found a participation

rate in the choice group as high as in the non-choice group.(49-51) A Dutch study offering FIT to non-participants of sigmoidoscopy showed that the overall attendance of sigmoidoscopy plus FIT was still lower than for FIT alone, while a similar Italian study showed the same attendance between those options.(51)In summary, based on the 56% higher CTC participation observed in a randomized controlled trial, CTC screening for colorectal cancer is more cost-effective than colonoscopy screening. The implementation of CTC screening requires prior satisfactory resolution of the optimal approach to managing extracolonic findings.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.
2. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. Cancer. 1986;58(2):397-403.
3. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol. 2008;103(6):1541-9.
4. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348(9040):1467-71.
5. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. J Natl Cancer Inst. 1999;91(5):434-7.
6. Shroff J, Thosani N, Batra S, Singh H, Guha S. Reduced incidence and mortality from colorectal cancer with flexible-sigmoidoscopy screening: a meta-analysis. World journal of gastroenterology. 2014;20(48):18466-76.
7. Nelson DB, McQuaid KR, Bond JH, Lieberman DA, Weiss DG, Johnston TK. Procedural success and complications of large-scale screening colonoscopy. Gastrointestinal endoscopy. 2002;55(3):307-14.
8. Liedenbaum MH, de Vries AH, Gouw CI, et al. CT colonography with minimal bowel preparation: evaluation of tagging quality, patient acceptance and diagnostic accuracy in two iodine-based preparation schemes. Eur Radiol. 2010;20(2):367-76.
9. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. The New England journal of medicine. 2007;357(14):1403-12.
10. de Haan MC, van Gelder RE, Graser A, Bipat S, Stoker J. Diagnostic value of CT-colonography as compared to colonoscopy in an asymptomatic screening population: a meta-analysis. Eur Radiol. 2011;21(8):1747-63.
11. Scott RG, Edwards JT, Fritschi L, Foster NM, Mendelson RM, Forbes GM. Community-based screening by colonoscopy or computed tomographic colonography in asymptomatic average-risk subjects. Am J Gastroenterol. 2004;99(6):1145-51.
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. Hanly P, Skally M, Fenlon H, Sharp L. Cost-effectiveness of computed tomography colonography in colorectal cancer screening: a systematic review. Int J Technol Assess Health Care. 2012;28(4):415-23.
14. Hassan C, Pickhardt PJ. Cost-effectiveness of CT colonography. Radiol Clin North Am. 2013;51(1):89-97.
15. Mavranezouli I, East JE, Taylor SA. CT colonography and cost-effectiveness. Eur Radiol. 2008;18(11):2485-97.
16. de Haan MC, Thomeer M, Stoker J, Dekker E, Kuipers EJ, van Ballegooijen M. Unit costs in population-based colorectal cancer screening using CT colonography performed in university hospitals in The Netherlands. Eur Radiol. 2013;23(4):897-907.
17. Stoop EM. Population-based colorectal cancer screening by colonoscopy or CT-colonography Rotterdam: Erasmus MC, 2013.

18. Lansdorp-Vogelaar I, van Ballegooijen M, Boer R, et al. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. Cancer 2009;115(11):2410-9.
19. Loeve F, Boer R, van Oortmarssen GJ, et al. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res 1999;32(1):13-33.
20. Statistics Netherlands. Statline; [cited 2012 14 Mar]; Available from: http://statline.cbs.nl/StatWeb/default.aspx.
21. CBO guideline. NVMDL; [cited 2014 25 Sep]; Available from: http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf.
22. van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol. 2006;101(2):343-50.
23. Pickhardt PJ, Hassan C, Halligan S, Marmo R. Colorectal cancer: CT colonography and colonoscopy for detection--systematic review and meta-analysis. Radiology. 2011;259(2):393-405.
24. Colquhoun P, Chen HC, Kim JI, et al. High compliance rates observed for follow up colonoscopy post polypectomy are achievable outside of clinical trials: efficacy of polypectomy is not reduced by low compliance for follow up. Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland. 2004;6(3):158-61.
25. 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.
26. Ness RM, Holmes AM, Klein R, Dittus R. Utility valuations for outcome states of colorectal cancer. Am J Gastroenterol. 1999;94(6):1650-7.
27. Leffler DA, Kheraj R, Garud S, et al. The incidence and cost of unexpected hospital use after scheduled outpatient endoscopy. Arch Intern Med. 2010;170(19):1752-7.
28. Dutch Health Care Authority. [website]: Law and regulations; [cited 2012 14 Mar]; Available from: http://www.nza.nl/regelgeving/tarieven/.
29. College voor zorgverzekeringen (CVZ); Farmacotherapeutisch Kompas 2014.
30. de Kok IM, Polder JJ, Habbema JD, et al. The impact of healthcare costs in the last year of life and in all life years gained on the cost-effectiveness of cancer screening. British journal of cancer. 2009;100(8):1240-4.
31. Yabroff KR, Lamont EB, Mariotto A, et al. Cost of care for elderly cancer patients in the United States. J Natl Cancer Inst. 2008;100(9):630-41.
32. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. PharmacoEconomics. 1997;11(2):159-68.
33. van den Berg M, de Wit GA, Vijgen SM, Busch MC, Schuit AJ. [Cost-effectiveness of prevention: opportunities for public health policy in the Netherlands]. Nederlands tijdschrift voor geneeskunde. 2008;152(23):1329-34.
34. Knudsen AB, Lansdorp-Vogelaar I, Rutter CM, et al. Cost-effectiveness of computed tomographic colonography screening for colorectal cancer in the medicare population. J Natl Cancer Inst. 2010;102(16):1238-52.
35. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Habbema JD. At what costs will screening with CT colonography be competitive? A cost-effectiveness approach. Int J Cancer. 2009;124(5):1161-8.

36. Denters MJ, Deutekom M, Bossuyt PM, Stroobants AK, Fockens P, Dekker E. Lower risk of advanced neoplasia among patients with a previous negative result from a fecal test for colorectal cancer. Gastroenterology. 2012;142(3):497-504.
37. Steele RJ, McClements PL, Libby G, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. Gut. 2009;58(4):530-5.
38. Weller D, Coleman D, Robertson R, et al. The UK colorectal cancer screening pilot: results of the second round of screening in England. British journal of cancer. 2007;97(12):1601-5.
39. Berrington de Gonzalez A, Kim KP, Knudsen AB, et al. Radiation-related cancer risks from CT colonography screening: a risk-benefit analysis. AJR American journal of roentgenology. 2011;196(4):816-23.
40. 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.
41. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bulletin of the World Health Organization. 2008;86(4):317-9.
42. Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375(9726):1624-33.
43. JE IJ, Tutein Nolthenius CJ, Kuipers EJ, et al. CT-Colonography vs. Colonoscopy for Detection of High-Risk Sessile Serrated Polyps. Am J Gastroenterol. 2016;111(4):516-22.
44. Sali L, Mascalchi M, Falchini M, et al. Reduced and Full-Preparation CT Colonography, Fecal Immunochemical Test, and Colonoscopy for Population Screening of Colorectal Cancer: A Randomized Trial. J Natl Cancer Inst. 2016;108(2).
45. Pyenson B, Pickhardt PJ, Sawhney TG, Berrios M. Medicare cost of colorectal cancer screening: CT colonography vs. optical colonoscopy. Abdominal imaging. 2015;40(8):2966-76.
46. Liedenbaum MH, Bipat S, Bossuyt PM, et al. Evaluation of a standardized CT colonography training program for novice readers. Radiology. 2011;258(2):477-87.
47. Plumb AA, Halligan S, Nickerson C, et al. Use of CT colonography in the English Bowel Cancer Screening Programme. Gut. 2014;63(6):964-73.
48. Hol L, Kuipers EJ, van Ballegooijen M, et al. Uptake of faecal immunochemical test screening among nonparticipants in a flexible sigmoidoscopy screening programme. Int J Cancer. 2012;130(9):2096-102.
49. Inadomi JM, Vijan S, Janz NK, et al. Adherence to colorectal cancer screening: a randomized clinical trial of competing strategies. Arch Intern Med. 2012;172(7):575-82.
50. Multicentre Australian Colorectal-neoplasia Screening G. A comparison of colorectal neoplasia screening tests: a multicentre community-based study of the impact of consumer choice. Med J Aust. 2006;184(11):546-50.
51. Segnan N, Senore C, Andreoni B, et al. Randomized trial of different screening strategies for colorectal cancer: patient response and detection rates. J Natl Cancer Inst. 2005;97(5):347 57

SUPPLEMENTARY MATERIAL

Appendix 1: **Extended methods**

The COCOS-trial

All colonoscopies were scheduled on one of the programmes of five experienced gastroenterologists (≥1000 colonoscopies. For bowel preparation, participants assigned to receive colonoscopy received 2 L of polyethylene electrolyte glycol solution (Moviprep, Norgine bv, Amsterdam, Netherlands) together with 2 L of transparent fluid, and a low-fibre diet for 2 days. For participants assigned to receive CT colonography, a non-cathartic preparation consisting of two times 50 mL of iodinated contrast agent (Telebrix Gastro, Guerbet, Aulnaysous-Bois, France) was given on the day before the examination, 50 mL 1·5 h before the examination, and a low-fibre diet for 1 day We obtained CT images in the supine and prone position on two 64-slice CT-scanners (Brilliance, Philips Healthcare, Best, Netherlands; SOMATOM Sensation, Siemens Medical Solutions, Erlangen, Germany) using a low-dose protocol. The attendance and yield of one screening round were compared. In summary, 5924 and 2920 people were invited for screening with colonoscopy and CTC, respectively, of which 1276 (21.5%) and 982 (33.6%) participated.(16) Participants with a lesion ≥10 mm at CTC were referred for diagnostic colonoscopy. All these patients attended colonoscopy. With this 10 mm cut-off, the detection rate of colorectal cancer was 0.5 for both tests per 100 individuals screened. The detection rate of large adenomas was 5.4 and 6.3 per 100 individuals screened with CTC and colonoscopy, respectively and the detection rate of small/medium sized adenomas was 1.4 and 23.3 (see Appendix 2).

Model input

Natural history

The microsimulation model was calibrated to the age-, stage-, and localization-specific incidence of CRC as observed in the Netherlands before the introduction of screening (between 1999 and 2003)(10) and the age-specific prevalence and multiplicity distribution of adenomas (any size, see below for size distribution) as observed in autopsy studies. Survival after clinical diagnosis is based on 1989-2003 survival data obtained from the Dutch Comprehensive Cancer Center.(15) The adenoma dwell-time and the preclinical duration of CRC were calibrated to the rates of interval and surveillance-detected cancers observed in randomized controlled trials evaluating screening using guaiac fecal occult blood tests and a once-only sigmoidoscopy and showed good concordance with the mortality reduction observed.(15)

Screening strategies

If no adenomas were detected during diagnostic colonoscopy after a positive CTC, the individual was assumed to be at low-risk for CRC and did not return to the screening program until after ten years. We assumed that surveillance colonoscopies would be performed until age 75 years or until the stop age for screening, whichever was latest.

Costs

Costs for a negative and positive screening CTC and colonoscopy were estimated using data from the COCOS trial.(18, 19) A positive CTC was estimated to cost €153, a negative CTC €144, a positive colonoscopy (including polypectomy and pathology costs) €329 and a negative colonoscopy €192. Invitation costs were estimated at €4.99, non-responder costs were based on the average costs of a non-responder (non-response at first invitation, reminder, after colonoscopy intake or for the colonoscopy itself). The costs for a diagnostic colonoscopy after a positive CTC were assumed to be the same as for primary colonoscopy. The costs of colonoscopy complications were obtained from a US cost-analysis of cases of unexpected hospital use after endoscopy in 2007 and multiplied with exchange rates for euros.(20) Costs for treatment of CRC were divided into three clinically relevant phases of care: initial treatment, continuous care and terminal care. Initial treatment costs were based on DTC-rates (Diagnosis Treatment Combination), except for oxaliplatin.(21) The costs for oxaliplatin were derived from the Dutch Health Care Insurance Board.(22) We assumed that during the continuous care phase, individuals would follow the Dutch CRC treatment guidelines,(23) and costs for periodic control were based on DTC rates. Terminal care costs were based on a Dutch last-year-of-life-cost-analysis.(24) We assumed that these costs increased with stage at diagnosis, at a rate observed for US patients.(25, 26) Dutch terminal care costs for individuals who died from CRC were approximately 40% of the US costs. We therefore assumed that terminal care costs of CRC patients who die from other causes were also 40% of the US costs. All costs were inflation adjusted to 2012 euro's using the Dutch Consumer Price Index.

Utility losses

We assumed an utility loss of 0.5 during 1.5 days resulting in 0.0021 QALYs lost per CTC and colonoscopy, and a loss of two weeks of life per complication (0.0384 QALYs). We also assigned a utility loss to each life-year with CRC care (Table 1).(27)

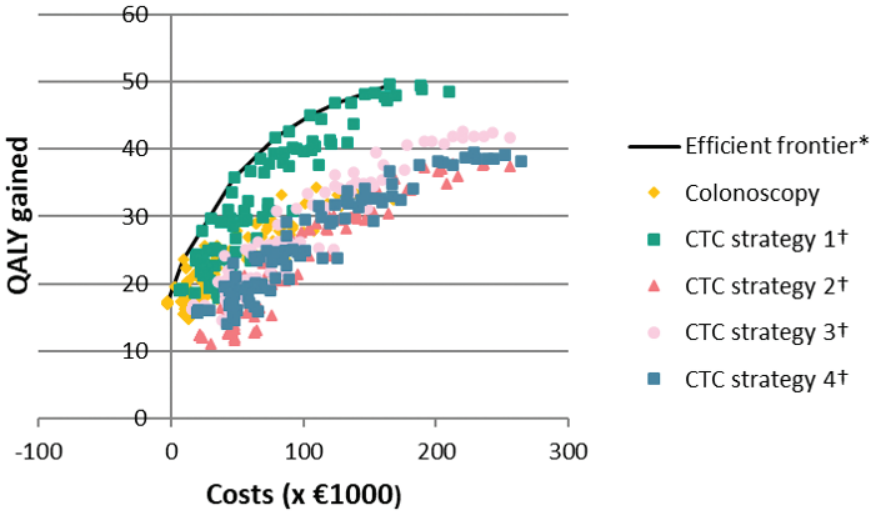
Appendix 2: **Detection rates in attenders to first round screening with colonoscopy and CTC with diagnostic colonoscopy as observed in the randomized controlled screening trial and as in the MISCAN-Colon model.**

	Observed	Simulated*
Colonoscopy		
DR small adenomas	23.3	21.6
DR large adenomas	6.3	6.7
DR colorectal carcinomas	0.5	0.8
CTC with 10 mm cut-off		
DR small adenomas	1.4	0.7
DR large adenomas	5.4	5.9
DR colorectal carcinomas	0.5	0.7
PR	8.6	9.2
Sent to CTC follow-up	8.4	10.6

*We validated the detection rates of colonoscopy with the test characteristics described in the main text. We then adjusted the size distribution of non-progressive adenomas to fit the observed detection rates best, without adjusting the overall prevalence of adenomas. Then, we validated the detection rates of CTC with the test characteristics described in the main text.

Appendix 3: **QALY's and costs in the sensitivity analyses**

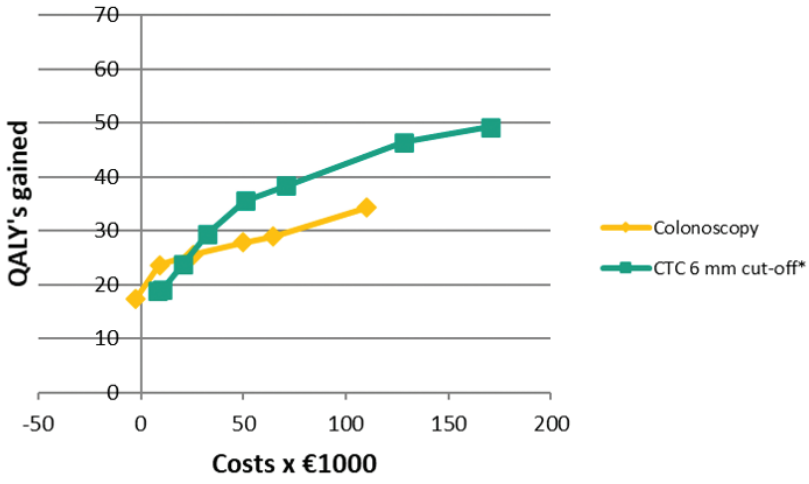
3a: Sensitivity analysis of all 480 screening strategies with observed attendance



	Screentest	Start age	Stop age	Interval	# Screens	QALY gained*†	Costs*† (€)	Costs/QALY gained†	ICER
	EFFICIENT STRATEGIES								
1	Colonoscopy	65	na	na	1	17	-3,013	-177	-177
2	Colonoscopy	60	na	na	1	17	-2,544	-146	1,182
3	Colonoscopy	55	65	10	2	24	9,208	390	1,883
4	CTC 6mm cut-off	55	70	5	4	36	47,667	1,332	3,162
5	CTC 6mm cut-off	55	70	3	6	42	78,877	1,891	5,258
6	CTC 6mm cut-off	50	70	3	7	45	104,569	2,326	7,916
7	CTC 6mm cut-off	50	75	3	9	47	123,725	2,644	10,398
8	CTC 6mm cut-off	45	80	3	12	50	165,129	3,328	14,709

3b: Sensitivity analysis including available data on extracolonic findings

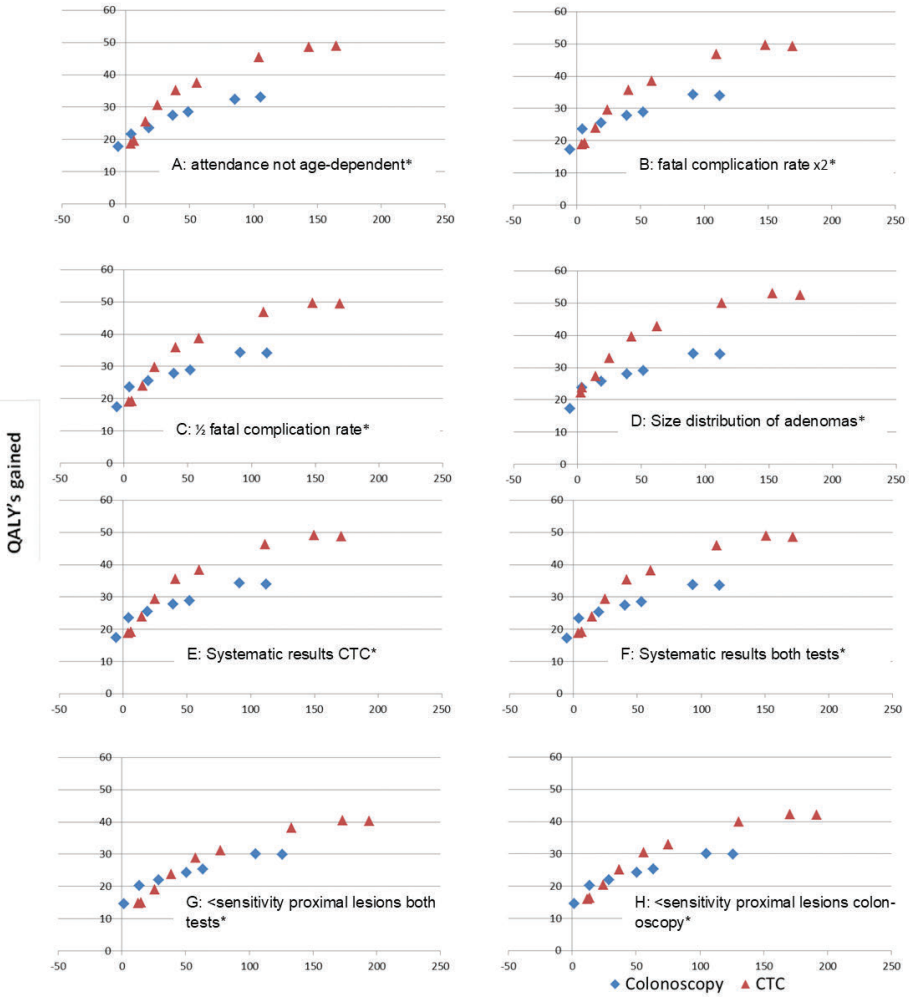
Scenario A*

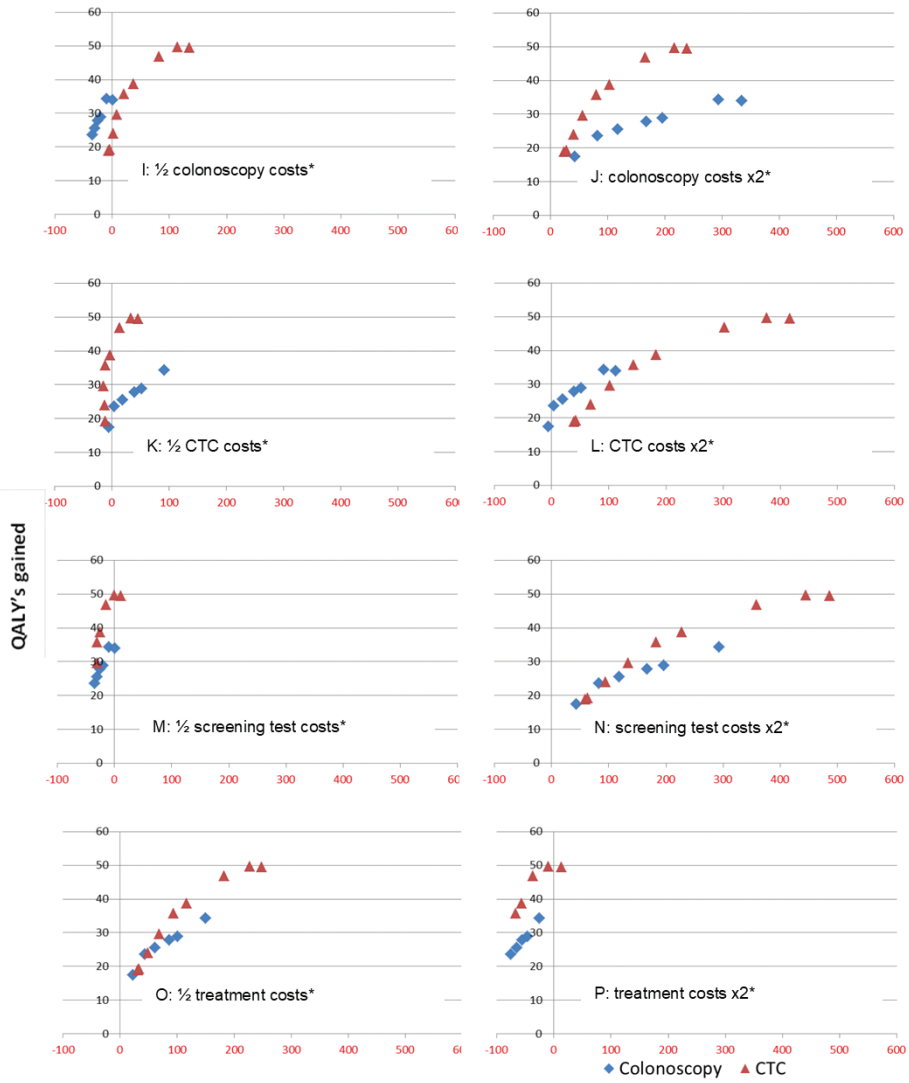


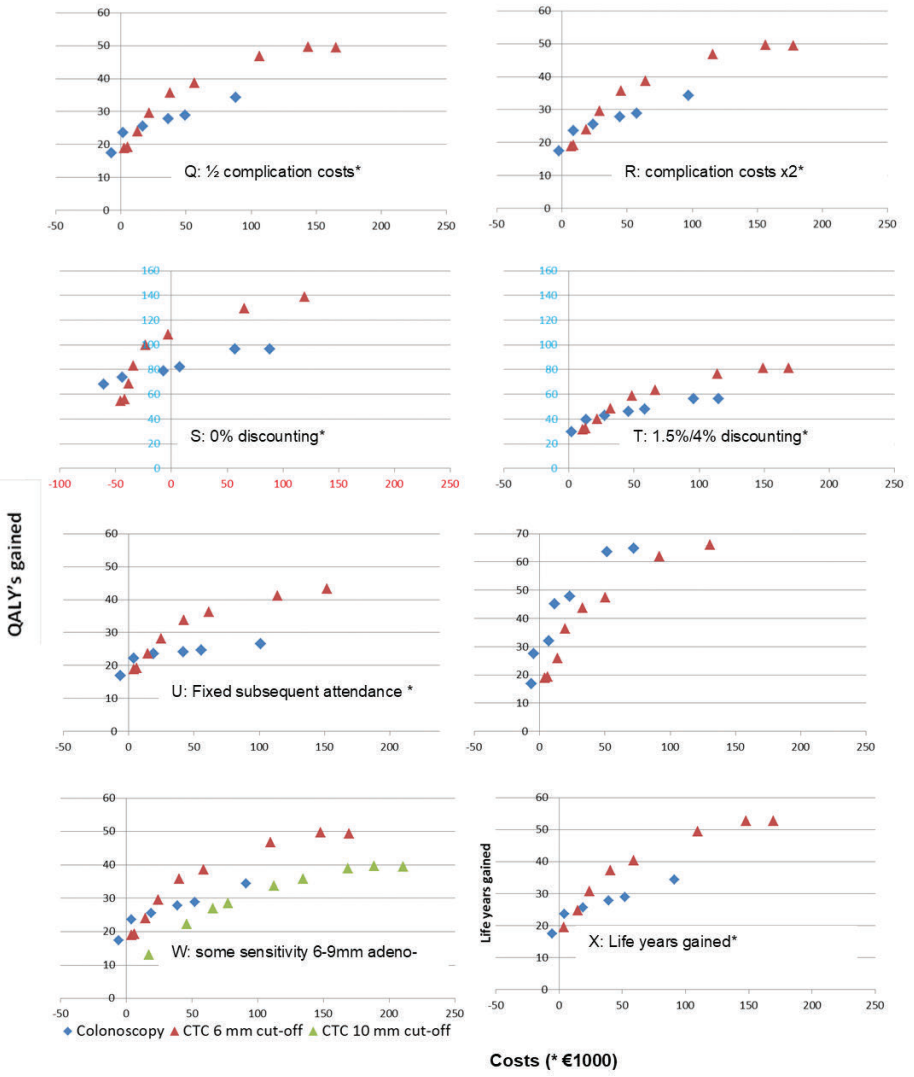
Scenario B*

* In scenario A we assumed only disadvantages of extracolonic findings: costs and utility loss due to extra procedures. In scenario B we also assumed a benefit of finding abdominal aortic aneurysms (AAA). We assumed a rate of extracolonic findings in the first screening as observed in the COCOS trial (9.6%),(1) and assumed the rate would drop with the same decrease as in Sconfienze et al.(2) to 2.6%. We assumed a follow up procedure due to detection of an extracolonic finding would cost €75.(3, 4) We assumed 2 days of life lost due to each follow up procedure. For AAA, we assumed the same detection rate as observed in the COCOS (0.7%); (1)that 10% of each detected AAA would prevent one death; and that 10 life years were gained per prevented death.(5)

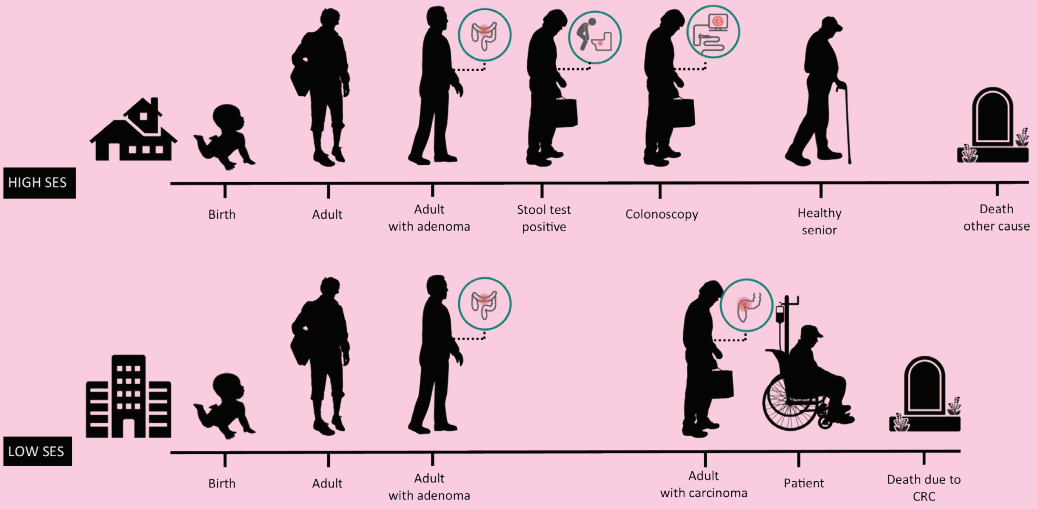
3c: QALY's and costs of efficient strategies in all other sensitivity analyses




I: ½ colonoscopy costs*
J: colonoscopy costs x2*
K: ½ CTC costs*
L: CTC costs x2*
M: ½ screening test costs*
N: screening test costs x2*
O: ½ treatment costs*
P: treatment costs x2*
QALY's gained
Costs (* €1000)
Colonoscopy
CTC



Sensitivity analyses with changes in input in A: attendance is not age-dependent; B: Double fatal complication rate; C: Half fatal complication rate; D: Size distribution of adenomas based on autopsy studies; E: Systematic false-negative testing in CTC; F: Systematic false-negative testing in CTC and colonoscopy; G: 50% less sensitivity for proximal lesions in colonoscopy and CTC; H: 50% less sensitivity for proximal lesions in colonoscopy; I: Half colonoscopy costs; J: Double colonoscopy costs; K: Half CTC costs; L: Double CTC costs; M: Half screening test costs; N: Double screening test costs; O: Half treatment costs; P: Double treatment costs; Q: Half complication costs; R: Double complication costs; S:0% discounting; T: 1.5% discounting on QALY's and 4% on costs; U: Fixed subsequent attendance; V: Random subsequent attendance; W: 10% of the sensitivity with 6 mm cut-off for medium sized adenomas in CTC with a 10 mm cut-off; X: Life years gained instead of QALYs gained


HIGH SES
Birth
Adult
Adult with adenoma
Stool test positive
Colonoscopy
Healthy senior
Death other cause
LOW SES
Birth
Adult
Adult with adenoma
Adult with carcinoma
Patient
Death due to CRC

Chapter 6

Socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal 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, M.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

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. 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. 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%, 95%CI 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%). 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 program, 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,[2, 3] aiming for equal access to CRC screening for the entire population.

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 organized screening programmes across the world have chosen to use FIT,[3] it is important to get more insight into the potential impact of a FIT screening programme on inequalities in health.

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.

The analyses were conducted with R-3.2.3.

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.

Data availability

The data that support the findings of this study are available from FSB. Restrictions apply to the availability of these data, which were used under license for this study. Data are available with the permission of FSB.

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.**

			Gender			Age	
	Number	**%**	**Males**	**%**		**median**	
Quintile 1	334233	17.9%	166013	49.7%		65.7	
Quintile 2	381344	20.4%	189929	49.8%		65.8	
Quintile 3	403907	21.6%	199777	49.5%		66.0	
Quintile 4	388664	20.8%	191341	49.2%		66.4	
Quintile 5	357912	19.2%	172222	48.1%		66.8	
Total	1866060	100.0%	919282	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%). Multivariate 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%) (OR Quintile 5 versus Quintile 1: 0.73 95% CI 0.69-0.77).

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

Quintile	N	Attendance to FIT	OR (univariate)	OR (multivariate)*	95% CI	
Quintile 1	246858	73.9%	1	1		p<0.001
Quintile 2	286527	75.1%	1.07	1.07	1.06 - 1.08	
Quintile 3	303133	75.1%	1.06	1.07	1.06 - 1.08	
Quintile 4	283640	73.0%	0.96	0.96	0.95 - 0.97	
Quintile 5	239945	67.0%	0.72	0.73	0.72 - 0.74	
	N	**Positivity rate**	**OR (univariate)**	**OR (multivariate)***	**95% CI**	
Quintile 1	14466	5.8%	1	1		p<0.001
Quintile 2	17726	6.2%	1.06	1.05	1.03 - 1.08	
Quintile 3	19235	6.3%	1.09	1.08	1.06 - 1.10	
Quintile 4	19037	6.7%	1.16	1.15	1.12 - 1.17	
Quintile 5	17145	7.1%	1.24	1.22	1.20 - 1.25	
	N	**Attendance to diagnostic colonoscopy**	**OR (univariate)**	**OR (multivariate)***	**95% CI**	
Quintile 1	11768	81.3%	1	1		p<0.001
Quintile 2	14612	82.4%	1.08	1.08	1.02 - 1.14	
Quintile 3	15732	81.8%	1.03	1.04	0.98 - 1.10	
Quintile 4	15234	80.0%	0.92	0.93	0.88 - 0.98	
Quintile 5	12992	75.8%	0.72	0.73	0.69 - 0.77	

* The multivariate OR is corrected for age and gender.

Diagnostic yield

The PPV for AN was highest in Quintile 3 (58.4%) and lowest in Quintile 5 (56.1%). Multivariate 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%, adjusted OR compared to Quintile 1 1.03 (95% CI 0.95-1.11)) and lowest in Quintile 4 (8.5%, adjusted OR compared to Quintile 1 0.90 (95% CI 0.82-0.97)) (Table 3).

The DR for AN in FIT participants was lowest in Quintile 1 (2.71% uncorrected and 3.33% corrected) and gradually increased with higher quintile (Quintile 5: 3.04% uncorrected, 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 3: **The positive predictive value (PPV) of FIT for advanced neoplasia (AN) and colorectal cancer (CRC) in each SES quintile, with the univariate and multivariate odds ratio (OR) and 95% CI.**

	N	PPV AN*	OR (univariate)	OR (multivariate)**	95% CI	
Quintile 1	6689	56.8%	1	1		p<0.001
Quintile 2	8388	57.4%	1.02	1.02	0.97 - 1.07	
Quintile 3	9191	58.4%	1.07	1.06	1.01 - 1.12	
Quintile 4	8872	58.2%	1.06	1.06	1.01 - 1.11	
Quintile 5	7295	56.1%	0.97	0.98	0.93 - 1.03	
	N	**PPV CRC***	**OR (univariate)**	**OR (multivariate)****	**95% CI**	
Quintile 1	1103	9.4%	1	1		p<0.01
Quintile 2	1376	9.4%	1.01	1.00	0.92 - 1.09	
Quintile 3	1516	9.6%	1.03	1.03	0.95 - 1.11	
Quintile 4	1301	8.5%	0.90	0.90	0.82 - 0.97	
Quintile 5	1165	9.0%	0.95	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 multivariate OR is corrected for age and gender.

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 multivariate odds ratio (OR) and 95% CI.**

		DR PER PARTICIPANT						YIELD PER INVITEE				
	N	DR AN uncorrected*	DR AN corrected**	OR (univariate)	OR (multi-variate)***	95% CI		yield AN	OR (univariate)	OR (multi-variate)***	95% CI	
Quintile 1	6689	2.71%	3.33%	1	1		p<0.01	2.00%	1	1		p<0.01
Quintile 2	8388	2.93%	3.55%	1.09	1.07	1.04 - 1.10		2.20%	1.10	1.10	1.06 - 1.13	
Quintile 3	9191	3.03%	3.70%	1.13	1.12	1.09 - 1.15		2.28%	1.14	1.14	1.10 - 1.17	
Quintile 4	8872	3.13%	3.91%	1.16	1.18	1.15 - 1.21		2.28%	1.14	1.13	1.10 - 1.17	
Quintile 5	7295	3.04%	4.01%	1.12	1.21	1.18 - 1.24		2.04%	1.01	1.00	0.97 - 1.04	
	N	DR CRC*	DR CRC**	OR (univariate)	OR (multi-variate)***	95% CI		yield CRC	OR (univariate)	OR (multi-variate)***	95% CI	
Quintile 1	1103	0.45%	0.55%	1	1		p<0.01	0.33%	1	1		p<0.01
Quintile 2	1376	0.48%	0.59%	1.08	1.06	1.01 - 1.12		0.36%	1.09	1.09	1.00 - 1.18	
Quintile 3	1516	0.50%	0.61%	1.12	1.12	1.07 - 1.17		0.38%	1.14	1.13	1.04 - 1.22	
Quintile 4	1301	0.46%	0.57%	1.03	1.05	0.99 - 1.10		0.33%	1.01	1.00	0.92 - 1.08	
Quintile 5	1165	0.49%	0.64%	1.09	1.17	1.12 - 1.23		0.33%	0.99	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 multivariate OR is corrected for age and gender and in the analysis per participant we corrected the DR for non-compliance to colonoscopy using poststratification (assuming full compliance).

Background CRC incidence

In total, 65,130 incident cases of CRC were recorded from 2008 to 2012. The European age-standardized 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).

Table 5: **The number of colorectal cancer cases recorded between 2008 and 2012 and the European age-standardized 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

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 (Appendix 1). 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 European Union (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 between 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]

A 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 Screen-IT, 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 organized 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[25] and sending an enhanced reminder letter with a banner that reiterates the screening offer.[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, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011;61(2):69-90.
2. Gupta S, Sussman DA, Doubeni CA, Anderson DS, Day L, Deshpande AR et al. Challenges and possible solutions to colorectal cancer screening for the underserved. Journal of the National Cancer Institute. 2014;106(4):dju032. doi:10.1093/jnci/dju032.
3. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49. doi:10.1136/gutjnl-2014-309086.
4. Aarts MJ, Lemmens VE, Louwman MW, Kunst AE, Coebergh JW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. European journal of cancer (Oxford, England : 1990). 2010;46(15):2681-95. doi:10.1016/j.ejca.2010.04.026.
5. de Klerk CM, Gupta S, Dekker E, Essink-Bot ML. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. Gut. 2017. doi:10.1136/gutjnl-2016-313311.
6. Javanparast S, Ward P, Young G, Wilson C, Carter S, Misan 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. doi:10.1016/j.ypmed.2010.02.003.
7. Mansouri D, McMillan DC, Grant Y, Crighton EM, Horgan PG. The impact of age, sex and socioeconomic deprivation on outcomes in a colorectal cancer screening programme. PloS one. 2013;8(6):e66063. doi:10.1371/journal.pone.0066063.
8. Pornet C, Dejardin O, Morlais F, Bouvier V, Launoy G. Socioeconomic determinants for compliance to colorectal cancer screening. A multilevel analysis. Journal of epidemiology and community health. 2010;64(4):318-24. doi:10.1136/jech.2008.081117.
9. Pruitt SL, Shim MJ, Mullen PD, Vernon SW, Amick BC, 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. doi:10.1158/1055-9965.epi-09-0135.
10. von Wagner C, Good A, Wright D, Rachet B, Obichere A, Bloom S 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. doi:10.1038/sj.bjc.6605392.
11. Wee LE, Koh GC, Chin RT, Yeo WX, Seow B, Chua D. 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. doi:10.1016/j.ypmed.2012.04.011.
12. Hurtado JL, Bacigalupe A, Calvo M, Esnaola S, Mendizabal N, Portillo I et al. Social inequalities in a population based colorectal cancer screening programme in the Basque Country. BMC public health. 2015;15:1021. doi:10.1186/s12889-015-2370-5.
13. Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I et al. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels. Gastroenterology. 2017;152(4):767-75.e2. doi:10.1053/j.gastro.2016.11.022.

14. Research TNIfS. Statusscores. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores. Accessed 11/09/2016 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. Accessed 05/10/2016 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. doi:10.1055/s-0032-1309788.
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, van Lenthe FJ, Coebergh JW, Janssen-Heijnen ML. 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. doi:10.1038/sj.bjc.6605949.
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, Spaander MCW, Dekker E, Koning HJ et al. Stage distribution of screen-detected colorectal cancers in the Netherlands. Gut. 2017. doi:10.1136/gutjnl-2017-315111.
22. Deutekom M, van Rijn AF, Dekker E, Blaauwgeers H, Stronks K, Fockens P 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. doi:10.1093/eurpub/ckp051.
23. Aarts MJ, Voogd AC, Duijm LE, Coebergh JW, Louwman WJ. 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. doi:10.1007/s10549-011-1363-z.
24. van Dam L, Korfage IJ, Kuipers EJ, Hol L, van Roon AH, Reijerink JC et al. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? European journal of cancer (Oxford, England : 1990). 2013;49(10):2321-30. doi:10.1016/j.ejca.2013.03.007.
25. Decker KM, Demers AA, Nugent Z, Biswanger N, Singh H. 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. doi:10.1136/bmjopen-2015-009470.
26. Wardle J, von Wagner C, Kralj-Hans I, Halloran SP, Smith SG, McGregor LM 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 (London, England). 2016;387(10020):751-9. doi:10.1016/s0140-6736(15)01154-x.

SUPPLEMENTARY MATERIAL

Sensitivity analyses of all outcomes with deciles instead of deciles

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

			Gender			Age	
	Number	%	Males	%		median	
Decile 1	152290	8.2%	75653	49.7%		65.7	
Decile 2	181943	9.8%	90360	49.7%		65.7	
Decile 3	187725	10.1%	93555	49.8%		65.7	
Decile 4	193619	10.4%	96374	49.8%		65.9	
Decile 5	201600	10.8%	99731	49.5%		66.0	
Decile 6	202307	10.8%	100046	49.5%		65.9	
Decile 7	200609	10.8%	99064	49.4%		65.9	
Decile 8	188055	10.1%	92277	49.1%		66.7	
Decile 9	189994	10.2%	91373	48.1%		66.8	
Decile 10	167918	9.0%	80849	48.1%		66.8	
Total	1866060	100.0%	919282	49.3%	p<0.001	65.9	p<0.001

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

Decile	N	Attendance to FIT	OR (univariate)	OR (multi-variate)*	95% CI	
Decile 1	110528	72.6%	1	1		p<0.0001
Decile 2	136330	74.9%	1.13	1.13	1.11 - 1.15	
Decile 3	141021	75.1%	1.14	1.14	1.13 - 1.16	
Decile 4	145506	75.2%	1.14	1.15	1.13 - 1.17	
Decile 5	151628	75.2%	1.15	1.15	1.14 - 1.17	
Decile 6	151505	74.9%	1.13	1.13	1.12 - 1.15	
Decile 7	147844	73.7%	1.06	1.07	1.05 - 1.08	
Decile 8	135796	72.2%	0.98	0.99	0.98 - 1.01	
Decile 9	131945	69.4%	0.86	0.87	0.86 - 0.88	
Decile 10	108000	64.3%	0.68	0.69	0.68 - 0.70	
	N	**Positivity rate**	**OR (univariate)**	**OR (multi-variate)***	**95% CI**	
Decile 1	6398	5.8%	1	1		p<0.0001
Decile 2	8068	5.9%	1.02	1.02	0.99 - 1.06	
Decile 3	8674	6.1%	1.07	1.06	1.03 - 1.10	
Decile 4	9052	6.2%	1.08	1.07	1.04 - 1.11	
Decile 5	9559	6.3%	1.10	1.09	1.05 - 1.12	
Decile 6	9676	6.4%	1.11	1.10	1.07 - 1.14	
Decile 7	9654	6.5%	1.14	1.13	1.09 - 1.16	
Decile 8	9383	6.9%	1.21	1.19	1.15 - 1.23	
Decile 9	9173	6.9%	1.22	1.20	1.16 - 1.24	
Decile 10	7972	7.3%	1.30	1.28	1.24 - 1.33	

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

	N	Attendance to diagnostic colonoscopy	OR (univariate)	OR (multi-variate)*	95% CI	
Decile 1	5155	80.6%	1	1		p<0.0001
Decile 2	6613	82.0%	1.10	1.10	1.01 - 1.19	
Decile 3	7172	82.7%	1.15	1.15	1.06 - 1.25	
Decile 4	7440	82.2%	1.11	1.12	1.03 - 1.21	
Decile 5	7862	82.2%	1.12	1.12	1.04 - 1.22	
Decile 6	7870	81.3%	1.05	1.06	0.98 - 1.15	
Decile 7	7830	81.1%	1.04	1.04	0.96 - 1.13	
Decile 8	7404	78.9%	0.90	0.91	0.84 - 0.99	
Decile 9	7068	77.1%	0.81	0.82	0.76 - 0.89	
Decile 10	5924	74.3%	0.70	0.71	0.65 - 0.77	

* The multivariate 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 decile, with the univariate and multivariate odds ratio (OR) and 95% CI.**

	N	PPV AN*	OR (univariate)	OR (multi-variate)**	95% CI	
Decile 1	2863	55.5%	1	1		p<0.001
Decile 2	3826	57.9%	1.10	1.10	1.02 - 1.18	
Decile 3	4125	57.5%	1.08	1.08	1.00 - 1.16	
Decile 4	4263	57.3%	1.07	1.07	1.00 - 1.15	
Decile 5	4613	58.7%	1.14	1.13	1.06 - 1.22	
Decile 6	4578	58.2%	1.11	1.11	1.03 - 1.19	
Decile 7	4608	58.9%	1.14	1.15	1.07 - 1.23	
Decile 8	4264	57.6%	1.09	1.09	1.01 - 1.17	
Decile 9	4024	56.9%	1.06	1.06	0.99 - 1.14	
Decile 10	3271	55.2%	0.99	1.00	0.92 - 1.07	
	N	**PPV CRC***	**OR (univariate)**	**OR (multi-variate)****	**95% CI**	
Decile 1	481	9.3%	1	1		p=0.04
Decile 2	622	9.4%	1.01	1.01	0.89 - 1.14	
Decile 3	673	9.4%	1.01	1.00	0.89 - 1.14	
Decile 4	703	9.4%	1.01	1.01	0.89 - 1.14	
Decile 5	759	9.7%	1.04	1.03	0.91 - 1.16	
Decile 6	757	9.6%	1.03	1.03	0.91 - 1.16	
Decile 7	683	8.7%	0.93	0.92	0.82 - 1.04	
Decile 8	618	8.3%	0.88	0.87	0.77 - 0.99	
Decile 9	655	9.3%	0.99	0.98	0.86 - 1.11	
Decile 10	510	8.6%	0.92	0.90	0.79 - 1.03	

*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 multivariate OR is corrected for age and gender.

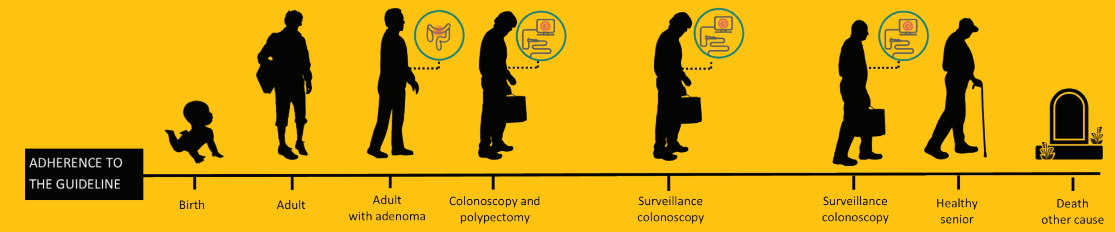
Table 4: **The detection rate (DR) per 100 participants uncorrected and corrected for colonoscopy uptake and the yield per 100 invitees of advanced neoplasia (AN) and colorectal cancer (CRC) for each decile, with the univariate and multivariate odds ratio (OR) and 95% CI.**

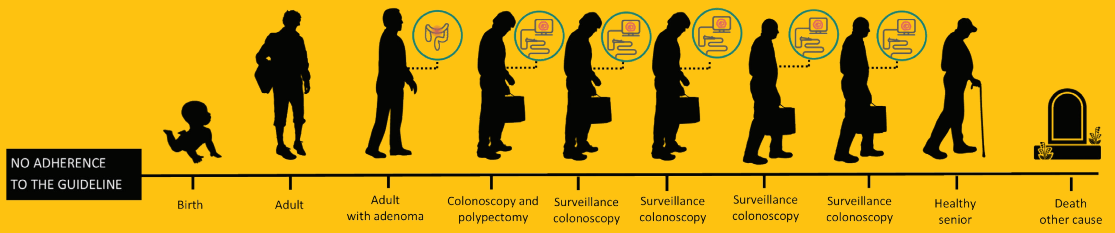
		DR PER PARTICIPANT						YIELD PER INVITEE				
	N	DR AN uncorrected*	DR AN corrected**	OR (univariate)	OR (multivariate)***	95% CI		yield AN	OR (univariate)	OR (multivariate)***	95% CI	
Decile 1	2863	2.59%	3.21%	1	1		p<0.01	1.88%	1	1		p<0.0001
Decile 2	3826	2.81%	3.42%	1.09	1.07	1.02 - 1.11		2.10%	1.12	1.12	1.07 - 1.17	
Decile 3	4125	2.93%	3.53%	1.14	1.10	1.06 - 1.15		2.20%	1.17	1.17	1.11 - 1.22	
Decile 4	4263	2.93%	3.56%	1.14	1.11	1.07 - 1.16		2.20%	1.18	1.17	1.12 - 1.23	
Decile 5	4613	3.04%	3.70%	1.18	1.16	1.12 - 1.20		2.29%	1.22	1.21	1.16 - 1.27	
Decile 6	4578	3.02%	3.71%	1.18	1.16	1.12 - 1.20		2.26%	1.21	1.21	1.15 - 1.26	
Decile 7	4608	3.12%	3.84%	1.21	1.20	1.16 - 1.25		2.30%	1.23	1.22	1.16 - 1.28	
Decile 8	4264	3.14%	3.98%	1.22	1.25	1.21 - 1.29		2.27%	1.21	1.19	1.14 - 1.25	
Decile 9	4024	3.05%	3.96%	1.18	1.24	1.20 - 1.29		2.12%	1.12	1.11	1.06 - 1.17	
Decile 10	3271	3.03%	4.08%	1.16	1.28	1.24 - 1.33		1.95%	1.02	1.01	0.96 - 1.06	
	N	DR CRC*	DR CRC**	OR (univariate)	OR (multivariate)***	95% CI		yield CRC	OR (univariate)	OR (multivariate)***	95% CI	
Decile 1	481	0.44%	0.54%	1	1		p<0.01	0.32%	1	1		p<0.001
Decile 2	622	0.46%	0.56%	1.05	1.03	0.95 - 1.11		0.34%	1.08	1.08	0.96 - 1.22	
Decile 3	673	0.48%	0.58%	1.10	1.07	0.99 - 1.15		0.36%	1.14	1.13	1.01 - 1.27	
Decile 4	703	0.48%	0.59%	1.11	1.09	1.01 - 1.16		0.36%	1.15	1.14	1.01 - 1.28	
Decile 5	759	0.50%	0.61%	1.15	1.13	1.05 - 1.20		0.38%	1.19	1.18	1.05 - 1.32	
Decile 6	757	0.50%	0.62%	1.15	1.14	1.07 - 1.21		0.37%	1.19	1.17	1.05 - 1.31	
Decile 7	683	0.46%	0.57%	1.06	1.05	0.98 - 1.13		0.34%	1.08	1.06	0.95 - 1.20	
Decile 8	618	0.46%	0.58%	1.05	1.07	0.99 - 1.15		0.33%	1.04	1.02	0.91 - 1.15	
Decile 9	655	0.50%	0.65%	1.14	1.19	1.12 - 1.27		0.34%	1.09	1.07	0.95 - 1.20	
Decile 10	510	0.47%	0.64%	1.09	1.18	1.10 - 1.26		0.30%	0.96	0.94	0.83 - 1.06	

*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 decile 1.

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


ADHERENCE TO THE GUIDELINE
Birth
Adult
Adult with adenoma
Colonoscopy and polypectomy
Surveillance colonoscopy
Surveillance colonoscopy
Healthy senior
Death other cause


NO ADHERENCE TO THE GUIDELINE
Birth
Adult
Adult with adenoma
Colonoscopy and polypectomy
Surveillance colonoscopy
Surveillance colonoscopy
Surveillance colonoscopy
Surveillance colonoscopy
Healthy senior
Death other cause

Chapter 7

Interpretation and adherence to the updated risk-stratified guideline for colonoscopy surveillance after polypectomy - a nationwide survey.

Miriam P. van der Meulen, Ida J. Korfage, Else-Mariëtte B. van Heijningen, Harry J. de Koning, Monique E. van Leerdam, Evelien Dekker, Iris Lansdorp-Vogelaar, on behalf of the working group on the guideline for colonoscopy surveillance.

in Press

ABSTRACT

Background: Low adherence to the Dutch guideline for colonoscopy surveillance after polypectomy led to the release of a new guideline in 2013. This new guideline was risk-stratified at a more detailed level than the previous one to achieve more efficient use of colonoscopy resources. This study assessed the feasibility of the risk-stratified guideline by evaluating the correct interpretation of and adherence to this guideline.

Methods: Based on semi-structured interviews with 10 gastroenterologists, we developed an online survey to evaluate gastroenterologists' recommendations for surveillance in 15 example cases of patients with polyps. If recommended intervals differed from the new guideline, respondents were asked to indicate their motives for doing so.

Results: Ninety-one out of 592 (15.4%) invited gastroenterologists responded to at least one case of whom 84 (14.2%) completed the survey. Gastroenterologists gave a correct recommendation in a median of 10 out of 15 cases, the adherence per case ranged from 14% to 95% (median case 76%). The two cases that addressed management of serrated polyps were least often answered correctly (14% and 28% correct answers). Discrepancies were mainly due to misinterpretation of the guideline with respect to serrated polyps (48%) or misreading of the questions (30%).

Conclusions: The median adherence to the updated colonoscopy surveillance guideline of 76% seems reasonable, and is higher than the adherence to the previous guideline (range: 22-80%, median 59%). This shows that detailed (more complex) risk stratification for designation of a surveillance interval is feasible. Adherence could potentially be improved by clarifying the correct interpretation on serrated polyps.

INTRODUCTION

Colorectal cancer (CRC) is the second most common cause of cancer mortality in the western world.[1] Individuals with adenomas are at increased risk of developing metachronous adenomas and CRC, even after the adenomas have been completely removed.[2-4] Therefore, colonoscopy surveillance after polypectomy is recommended.[5,6] The frequency of colonoscopy surveillance and the adherence to surveillance recommendations are important, since too little surveillance has the risk of diminishing the preventive effect of the surveillance program for CRC, while too intensive surveillance exposes the patient to unnecessary risks and burden and waste of colonoscopy as well as financial resources.

Colonoscopy is a scarce resource and many countries face waiting lists for these procedures.[7,8] With the implementation and expansion of CRC screening programs throughout the world,[9] the demand for colonoscopies will further increase.

Before the introduction of mass screening, colonoscopies for surveillance after polypectomy encompassed about 13% of all colonoscopies conducted in the Netherlands.[10] The recently started CRC screening program will result in an increase in adenoma diagnoses, eventually resulting in an increasing number of patients that meet the criteria for surveillance colonoscopy. This emphasizes the importance of efficient use of colonoscopy capacity and thus also of efficient surveillance strategies.

However, the colonoscopy capacity is often not used efficiently for surveillance. Current international guidelines only consider presence or absence of risk factors for metachronous advanced neoplasia, but do not take into account combinations of risk factors. Several surveys showed suboptimal adherence to guidelines for surveillance after polypectomy in daily practice, with clinicians often recommending too short surveillance intervals.[11-13] A Dutch study reported on 6 example cases that were assigned correct recommendations ranging from 22-80% (median 59%). In most of the incorrect recommendations, gastroenterologists used shorter surveillance intervals than prescribed by the national guideline.[12] This was caused by clinicians often incorporating other adenoma characteristics, like histology and size of the adenomas into their recommendation, even though at that time the Dutch surveillance guidelines only differentiated the recommended surveillance interval by adenoma multiplicity. [12]

The updated risk-stratified guideline for colonoscopy surveillance introduced in 2013 incorporated multiplicity, size, location and histology of adenomas as well as presence of large serrated lesions.[14] Through a score chart these polyp characteristics are combined into a risk score (0 - 5) to optimize the risk stratification of patients for designation of a surveillance interval. However, this new guideline is more complex than the previous guideline and most international guidelines. This may cause gastroenterologists to misunderstand or misinterpret the guideline, or potentially even

not use it all, eventually resulting in low adherence to the recommendations. Therefore, the aim of our study was to evaluate gastroenterologists' interpretation and adherence to this new guideline.

MATERIALS AND METHODS

Design

To assess the correct interpretation of and adherence to the Dutch guideline for colonoscopy surveillance after polypectomy, we developed an online survey consisting of 15 example cases of patients that underwent colonoscopy with polypectomy. The survey was pilot-tested during semi-structured interviews with 10 gastroenterologists. We sent the survey to all gastroenterologists in the Netherlands and asked them to designate their surveillance recommendation for each case. If recommendation(s) differed from the new guideline, their motives for doing so were asked for a maximum of 2 random example cases. It is estimated that the survey would take approximately 15 minutes and that information was provided to the gastroenterologists.

Dutch guideline for colonoscopy surveillance after polypectomy

The new Dutch guideline for surveillance after polypectomy was introduced in 2013. [14] The surveillance interval is based on the number of adenomas and the presence of at least one large adenoma (≥10mm), at least one villous adenoma (>75% villous component) and/or at least one proximal adenoma. Serrated polyps (including hyperplastic polyps, sessile serrated adenomas/polyps and traditional serrated adenoma) are incorporated in the guideline only if at least one serrated polyp measures ≥10mm. Other characteristics (total number, localisation) of the serrated polyps are not taken into account. High-grade dysplasia (HGD) in adenomas is not incorporated as a risk factor in the guideline as it is not confirmed to be an independent risk factor, probably because HGD is highly associated with other factors such as size. Using a score chart, the polyp characteristics are combined into a risk score (0 - 5), see Figure 1. The total risk score indicates a recommended surveillance interval of 3 or 5 years, or no surveillance at all.

Figure 1: **score chart of the Dutch guideline for colonoscopy surveillance after polypectomy.[14]**

SCORE TABLE FOR PRESENCE OF ADENOMA CHARACTERISTICS AND SERRATED POLYPS*		
Polyp Characteristics	Values	Points
Number of adenomas	1	0
	2-4	1
	≥5	2
Presence of at least one adenoma ≥10mm and/or one large serrated polyp ≥10mm**	No	0
	Yes	1
Presence of at least one villous adenoma***	No	0
	Yes	1
Presence of at least one proximal adenoma****	No	0
	Yes	1
Total risk score		

* A patient with 5 proximal serrated polyps of which 2 ≥ 10 mm fulfil the WHO criteria of the serrated polyposis syndrom; see the guideline of hereditary colorectal cancer.

**A serrated polyp encompasses: hyperplastic polyps, sessile serrated polyps/adenomas and traditional serrated adenomas

***An adenoma with at least 75% villous histology.

****Proximal is defined as cecum, colon ascendens, colon transversum and flexura lienalis

SURVEILLANCE INTERVAL BASED ON THE ADENOMA RISK SCORE	
Score during index colonoscopy	Interval after index colonoscopy
0	No surveillance*
1-2	5 years
3-5	3 years
Score during subsequent colonoscopy	Interval after subsequent colonoscopy
0	5 years**
1-2	5 years
3-5	3 years

*Patients with a score of 0 during index colonoscopy are advised to not undergo surveillance colonoscopy. These patient are sent back to the national screening programme in 10 years if aged 55-75 years at that moment.

**For patients in which a high-risk adenoma (score ≥3) was never detected, surveillance can be ended after two subsequent negative colonoscopies. These patient are sent back to the national screening programme in 10 years if aged 55-75 years at that moment.

Stopping age of surveillance: 75 years, unless the wish and condition of the patient justify a different stopping age

Survey

The survey consisted of three parts. The first part (baseline questions) contained 7 questions on (demographic) characteristics of the gastroenterologist: gender; age; type of hospital; specialisation; number of colonoscopy procedures per year; years of experience and if they perform colonoscopies for the national screening programme.

The second part consisted of 15 example cases of patients that underwent colonoscopy with polypectomy. To avoid bias and disadvantages for the later example cases if respondents would not finish the complete survey, there were two versions of the survey that only differed regarding the order of the example cases. The example

cases varied in age, gender, adenoma/polyp number, size and location of adenomas, grade of dysplasia and presence of (tubulo)villous histology, see Table 3 and Appendix 2. Respondents were informed that unless noted otherwise, all patients were in good health; had no familial risk for colorectal cancer; had undergone their first colonoscopy; bowel preparation was good; the cecum was reached; and the polyp was removed in one piece and endoscopically complete.

In each case, the gastroenterologist was asked to recommend the surveillance interval. Response options were: an interval of <1 to 10 years; no surveillance; surveillance only if the patient would be in good condition (at a 3 or 5-year interval); and referral to the clinical geneticist (Appendix 2).

In the third part of the survey respondents were given feedback on the recommendations they had given in part 2. For each case in which the recommendation did not meet the guideline, the respondent was shown a table with the interval they recommended versus the guideline-recommendation. Subsequently, the motives for deviation were asked for a maximum of two random example cases. Response options were: thinking the answer was in agreement with the guideline; not having read the question correctly; not familiar with the new guideline; based on scientific evidence or clinical experience; or an answer in the free text field (Appendix 2).

Pilot-tests

Interviews

10 gastroenterologists were interviewed between May and July 2014 (Appendix 1). The selected gastroenterologists differed in age, gender, setting (regional or academic hospital) and region. One of the authors (MvdM) conducted all interviews, which were audio-recorded. The interviews were semi-structured, starting with open questions on what gastroenterologists considered advantages and bottlenecks of the guideline. Then, they were presented 5 cases and were asked what interval they would recommend and why. Based on the response of the interviewed gastroenterologists the cases were improved and several answering options on why people would potentially deviate from the current guideline were added.

Online pilot

After enhancement of the survey due to the findings of the interviews, the survey was additionally validated by five medical researchers in gastroenterology of the Academic Medical Center (AMC) and the Netherlands Cancer Institute.

Distribution of the survey

The online survey was send by email to all 594 registered gastroenterologists of the Dutch Gastroenterology association in December 2014. A reminder of the survey was sent 6 weeks later in January 2015. The survey was anonymous and written in Dutch.

Statistical analyses

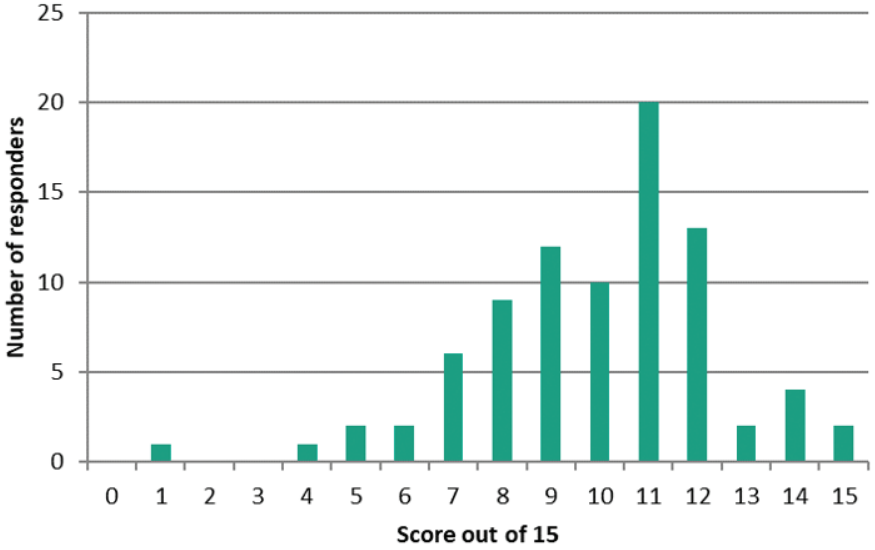
Statistical analyses were conducted with SPSS version 22.0 (IBM corporation, USA). To be considered as a respondent at least 4 baseline questions had to be answered. Descriptive statistics were used to analyze the data; medians and interquartile range (IQR) were calculated for non-normally distributed data. Outcomes were the number of respondents, the median number of correct recommendations per respondent – for those who responded to all cases -, and the number of correct recommendations per case. Differences between subgroups in correct recommendations per respondent were tested with the Mann-Whitney U test.

RESULTS

Of 592 invitees, 91 (15.4%) responded to at least 1 case. One respondent was excluded as he or she did not actively perform colonoscopies. Of the 91 responders, 84 (14.2% of 592 invitees) gastroenterologists responded to all cases.

Sixty-five percent of the respondents were male, the median age was 43 years old (Table 1). Most respondents worked in a hospital without gastroenterology trainees (43%), most had 0-10 years of experience (51%), performed more than 300 colonoscopies per year (70%) and performed colonoscopies for the national bowel cancer screening program (63%). Thirty-six percent of the respondents indicated that they did not consult the guideline during the questionnaire, while 48% used the pocket card of the guideline and 10% the app.

Figure 2: **Distribution of the score out of 15 example cases (number of correct answers according to the guideline) of the respondents to all example cases (n=84).**



The 84 respondents that indicated recommendations for all cases were correct in a median of 10 (out of 15) cases (IQR 8-11) (Table 2 and Figure 2). The number of correct recommendations did not differ by gender, age, type of hospital and participation in the screening program, but consulting the guideline during the questionnaire was associated with an increase in adherence (p=0.015).

The cases received a correct recommendation ranging from 14% to 95% per case (median case 76%) (Table 3). For all cases combined, a mean of 66% recommendation were correct, 22% of the recommended intervals were shorter than the guideline, 3% of the given recommended intervals were longer than the guideline, 7% gave no surveillance interval, but an alternative recommendation while a surveillance interval was recommended (such as referral to a clinical geneticist, or only referral if the patients was in good condition) and 2% recommended no surveillance at all while the guideline did recommend surveillance. In 48% of the discrepant cases, gastroenterologists were convinced they had recommended the correct interval, while in 30% of the discrepant cases, gastroenterologists had not read the question correctly (Table 4).

Table 1: **Baseline characteristics of the respondents (N = 91).**

Variable	N	
Age (median)	99	43 (IQR 35-52)
Gender		
Males	64	65%
Females	35	35%
Type of hospital		
Academic	19	19%
Non-academical teaching hospital	37	38%
Peripheral hospital	42	43%
Missing	*1*	
Specialisation		
Gastroenterologist	92	95%
Fellow	5	5%
Missing	*2*	
Years of experience		
None	2	2%
0 to 10	50	51%
10 to 20	21	21%
20 to 30	18	18%
30 to 40	6	6%
>40	2	2%
Colonoscopies per year		
< 150	10	10%
150-300	19	20%
>300	68	70%
Missing	*2*	
Performing colonoscopies for the screening programme		
Yes	61	63%
No	36	37%
Missing	*2*	
Use of source during questionnaire		
None	29	36%
App	8	10%
Pocket card	39	48%
Website	2	2%
2 sources	3	4%
Missing	*18*	

Table 2: **Score (median correct recommendations according to the guideline) out of 15 example cases of respondents to all example cases (n=84).**

		N	Score out of 15 cases	P-value
Gender	*Men*	51	10	
	Women	33	10	0.81
Age	*<40*	37	11	
	>40	47	10	0.62
Academic hospital	*Yes*	16	11	
	No	67	10	0.44
Performing colonoscopies for the CRC screening programme*	*Yes*	51	11	
	No	31	10	0.71
Use of source*	*Yes*	29	11	
	No	52	9	0.02
Total			10	

‡ Either use of no source at all, or use of the app, pocket card and/or website.

The recommendation for surveillance was least often correct for the cases on serrated lesions (case 10, 14% correct, and case 11, 28% correct) (Table 3). All discrepant answers recommended a shorter interval (86% and 72%) of which 92% and 95% recommended the interval that would be correct if serrated polyps would be scored the same as conventional adenomas. In 78% and 65%, respectively, of these discrepant cases, gastroenterologists had the impression they had recommended the correct interval. 13% and 26% respectively answered that they had not read the question correctly (Table 4).

Next, cases with a patient with older age (³75 years) were least often answered correctly, at 31% for case 8 and 52% for case 7 (Table 3). In the case of a 75-year old male with four adenomas and one adenoma with HGD (case 8), 40% of the respondents recommended a shorter interval than the guideline and 25% of the respondents recommended surveillance after five years. Responders motivated their discrepancy with the guideline for these cases because they were either convinced their answer was in accordance with the guideline or they had not read the question correctly (Table 4). Of those who provided an answer for case 7 and 8 in the free text field, 12 out of 14 mentioned they did not incorporate age or the condition of the patient at older age in their answer. In the case of a 79-year old male with five adenomas (case 7), the correct answer would be to recommend no surveillance, unless the patient remains in good condition, then in 3 years. Eleven percent of the respondents would not recommend any surveillance regardless of physical condition, and 26% of the respondents recommended surveillance after three years. If you assume that after these 3 years everyone would examine these older patients if they are still in good condition, 78% of cases would be answered correctly.

The case with a large tubulovillous adenoma (case 3) was correctly answered by only half (52%) of the gastroenterologists. If incorrect, recommended intervals were almost always too short, see Table 3. Discrepancies were again mainly due to misinterpretation of the guideline (62%). Three out of four answers in the free text field explained that they scored the tubulovillous adenoma equal to villous adenoma.

Remarkable about the case of the 65-years old male with one adenoma with HGD (case 9) was that even though 76% of the respondents answered correctly, the incorrect answers had a large discrepancy with the interval recommended by the guideline. Eleven percent of the respondents recommended a surveillance colonoscopy within one year, whereas a five-year interval is recommended by the guideline. Two out of six gastroenterologists that motivated their discrepancy from the guideline for this case responded that they consider lesions with HGD as high risk.

A new aspect in the guideline is that no surveillance is indicated if patients have only one distal non-advanced adenoma (case 1). This was correctly recommended by 84% of the respondents.

The remaining 8 cases were correctly answered by a median of 86% (58% to 95% per case) of the respondents.

Table 3: **Short description of the 15 example cases with the recommended interval and the results per example case.**

	DESCRIPTION CASE								RESULTS PER CASE					
	Common cases													
	Age	**G**	**# AD**	**Size (mm)**	**Vill.**	**HGD**	**# prox**	**Recommended interval**	**N**	**%corr**	**%early**	**%late**	**%no surv**	**%other**
1	60	M	1	8	T	no	0	No surveillance	86	84%	**16%**	0%	na	0%
2	69	M	1	12	T	no	0	5y	89	91%	1%	1%	3%	3%
3	54	M	1	20*	TV	no	1	5y	85	**52%**	**47%**	1%	0%	0%
4	62	F	2	22*	V	no	0	3y	84	79%	6%	**15%**	0%	0%
5	63	F	4	9	V	no	2	3y	84	90%	0%	7%	0%	2%
6	60	F	5	12	T	no	4	3y	84	95%	2%	1%	0%	1%
7	79	M	5	8	T	no	3	Only if healthy, then 3y[†]	84	**52%**	1%	8%	11%	**27%**
8	75	M	4	12	T	yes	0	Only if healthy, then 5y[†]	84	**31%**	**40%**	0%	4%	**25%**
9	65	M	1	11	TV	yes	0	5y	88	76%	**17%[§]**	1%	2%	3%
	Serrated adenomas/polyps													
	Age	**G**	**# SP**	**Size (mm)**			**# prox**		**N**	**% corr**	**% early**	**% late**	**%no surv**	**% other**
10	58	F	1	8			1	No surveillance	85	**14%**	**86%**	0%	na	0%
11	54	F	2	12			2	5y	86	**28%**	**72%**	0%	0%	0%
				Family history										
	Age	**G**	**Score**	**FM**	**Age FM**	**Previous examination**			**N**	**% corr**	**% early**	**% late**	**%no surv**	**% other**
12	51	M	2	Brother	<50	Yes, no hereditary CRC		5y	84	83%	14%	0%	0%	2%
13	53	M	1	Sister	<50	no		Refer to geneticist	88	**58%**	0%	0%	0%	**42%**
	Negative colonoscopies													
	Age	**G**	**Initial Score**	**# neg. colo**					**N**	**% corr**	**% early**	**% late**	**%no surv**	**% other**
14	69	M	4[‡]	1				5y	86	88%	5%	2%	3%	1%
15	63	F	2[‡]	2				No surveillance	86	73%	**23%**	0%	na	3%
	TOTAL								1283	66%	22%	3%	2%	7%

G = gender; # AD = the number of adenomas; Size = size of the largest lesion Vill. = presence of villousness, with T = tubular adenoma, TV = tubulovillous adenoma, V = villous adenoma; HGD = presence of high-grade dysplasia; # prox = the number of proximal adenomas; #SP = number of serrated polyps; FM = family member with CRC diagnoses; # neg. colo = number of previous negative colonoscopies; %corr = % of answers correct, according to the guideline (**red** ≤50% correct, **orange** ≤ 70% correct) %early = % of answers with a shorter interval than recommended; %late = % of answers with a longer interval than recommended; % no surv = % of answers with no surveillance while this is not recommended; %other = % of answers with another answer (**red** ≥40% of respondents, **orange** ≥15% of respondents)

* In the cases with adenomas ≥ 20 mm we describe that patients had had another colonoscopy after 6 months at which no residual tissue was found.

[†] If an individual will be 75 at the subsequent screening, then surveillance should only take place if the individual is still healthy, and an interval is based on the adenoma risk score. We defined all answers containing a shorter interval then recommend based on the adenoma risk score as “early”, we defined all answers containing a longer interval then recommend based on the adenoma risk score as “late” and an answer with the same interval but without the addition that the patient should only be screened if healthy as “other”.

[‡] Full findings at the initial colonoscopy were: 69 year old male: 2 adenomas: Polyp A was a distal villous adenoma of 12 mm. Polyp B was a proximal tubular adenoma of 8 mm with low-grade dysplasia; 63 year old female: 2 adenomas: Polyp A was a distal tubular adenoma of 5 mm with low-grade dysplasia. Polyp B was a distal tubular adenoma of 12 mm with low-grade dysplasia.

[§] 10 out of 15 of the respondents with an answer with a too short interval, answered they would offer a surveillance colonoscopy within a year.

Table 4: **Short description of the 15 example cases with the recommended interval and the motivation to deviate from the guideline**

	DESCRIPTION CASE								MOTIVATION TO DEVIATE FROM GUIDELINE				
	Common cases												
	Age	**G**	**# AD**	**Size (mm)**	**Vill.**	**HGD**	**# prox**	**Recommended interval**	**N**	**% expected to be correct**	**% based on clinical experience**	**% did not read correctly**	**% other**
1	60	M	1	8	T	no	0	No surveillance	6	0%	17%	83%	0%
2	69	M	1	12	T	no	0	5y	0				
3	54	M	1	20	TV	no	1	5y	21	62%	10%	10%	19%*
4	62	F	2	22	V	no	0	3y	6	50%	17%	33%	0%
5	63	F	4	9	V	no	2	3y	4	50%	0%	50%	0%
6	60	F	5	12	T	no	4	3y	1	0%	0%	100%	0%
7	79	M	5	8	T	no	3	Only if healthy, then 3y	16	25%	0%	31%	44%[ł]
8	75	M	4	12	T	yes	0	Only if healthy, then 5y	25	32%	0%	44%	24%[ł]
9	65	M	1	11	TV	yes	0	5y	6	50%	0%	17%	33%[‡]
	Serrated adenomas/polyps												
	Age	**G**	**# SP**	**Size (mm)**			**# prox**		**N**				
10	58	F	1	8			1	No surveillance	23	78%	0%	13%	9%
11	54	F	2	12			2	5y	23	65%	4%	26%	4%
	Family history												
	Age	**G**	**Score**	**FM**	**Age FM**	**Previous examination**			**N**				
12	51	M	2	Brother	<50	Yes, no hereditary CRC		5y	6	17%	17%	50%	17%
13	53	M	1	Sister	<50	no		Refer to geneticist	11	18%	9%	36%	36%
	Negative colonoscopies												
	Age	**G**	**Initial Score**	**# neg. colo**					**N**				
14	69	M	4	1				5y	3	67%	0%	33%	0%
15	63	F	2	2				No surveillance	6	83%	0%	17%	0%
	TOTAL								157	48%	4%	30%	17%

G = gender; # AD = the number of adenomas; Size = size of the largest lesion Vill. = presence of villousness, with T = tubular adenoma, TV = tubulovillous adenoma, V = villous adenoma; HGD = presence of high-grade dysplasia; # prox = the number of proximal adenomas; #SP = number of serrated polyps; FM = family member with CRC diagnoses; # neg. colo = number of previous negative colonoscopies; N = number of answers per case.

* 3 out of 4 respondents answered that they scored the tubulovillous adenoma as a villous adenoma.

[ł] 12 out of 14 other answers incorporated the age of the patient in their answer.

[‡] Both (2) respondents mentioned they saw HGD as high risk.

DISCUSSION

Using a survey with 15 example cases, we showed that the cases were assigned a recommend surveillance intervals that is in agreement with the current guideline in 14% to 95% per case (median case 76%) and the gastroenterologist gave a correct recommendation in a median of 10 cases. Cases involving serrated polyps or elderly patients were most often answered incorrectly.

As large inter- and intra-observer variation exists in pathologists for diagnosis various types of serrated polyps, serrated polyps are treated as one histological entity in the guideline. To prevent that patients with only small hyperplastic polyps will receive a surveillance recommendation, number and location of serrated polyps does not impact the length of the surveillance interval in the guideline.[14] In our survey almost all discrepant recommendations would have been correct if serrated polyps would be scored the same way as conventional adenomas. We therefore recommend to provide further clarification of the guideline on how to deal with serrated polyps. This could potentially be accompanied by further teaching sessions, for example an e-learning course for gastroenterologists is already implemented.

Prior to developing the survey, we hypothesized three other instances where gastroenterologists might deviate from the guideline: in cases with adenomas with high-grade dysplasia, in cases with tubulovillous adenomas, and in cases where the guideline recommends returning to the national CRC screening program with FIT. Although cases involving HGD were answered according to the guideline by a majority of respondents, the gastroenterologists that did not answer in line with the guideline recommended an interval shorter than one year. In the surveillance guideline of the US and in the guideline of the European Society of Gastrointestinal Endoscopy**,** HGD is considered a high-risk feature.[6,15] However, in the Dutch guideline HGD is not incorporated as a separate risk factor, because a meta-analysis and the study on which the guideline was based on did not confirm HGD as an independent risk factor in addition to the other factors.[4,16] This is mainly explained by the fact that HGD is rarely seen in small (<10mm) tubular or tubulovillous adenoma. Furthermore, there is a large interobserver variation between pathologists making this feature an unreliable risk factor. In the interviews, half of the gastroenterologists mentioned that they were not entirely convinced that HGD should not be incorporated, while one gastroenterologist in the interview specifically mentioned that HGD was not incorporated in this score chart, but should be considered as high-risk assigning a surveillance interval within one year.

Discrepancies for cases with a (tubulo)villous adenoma seem to be caused by gastroenterologists scoring tubulovillous adenomas as villous adenomas. However, in previous studies a tubulovillous adenoma (>25% and < 75% villous component) was not a risk factor for metachronous disease in a multivariable model,[4,17]. Only villous

adenoma (>75% villous component) was found to be a risk factor,[16] and therefore assigned an extra point to the risk score chart in the guideline. This might, however, be confusing because internationally an advanced adenoma is defined as an adenoma ≥10 mm, HGD, or a tubulovillous component (>25%). Also, during the interviews, 6 out of 10 gastroenterologists mentioned that adhering to the guideline was difficult considering the difference between tubulovillous and villous adenomas, because pathology reports in their hospital do not include percentages nor whether adenomas were villous or tubulovillous (Appendix 1).

In contrast to the cases discussed before, the case in which a person with only one distal non-advanced adenoma should return to the screening program was answered correctly by a large majority without striking discrepancies. Previously these patients would be recommended surveillance after six years, but apparently the change to recommend no surveillance is well accepted.

Adherence to our colonoscopy surveillance guideline is at the high end of adherence as reported in other studies. The median adherence to the guideline was reported to be 49% in France, 63% in Canada and 52.5% and 69% in the USA in two different periods. A study in the USA in primary care physicians found a far lower adherence of 29%.[11-13,18-20]. More specifically, compared to the reported adherence in the Netherlands when the simple 2002 guideline was implemented, our estimate of adherence shows a clear increase with a median of 76% adherence, compared to[12] a median of 59% (range: 22-80%) in the survey based on the old guideline. This comparison clearly indicates that more complex guidelines do not necessarily lead to confusion and lower adherence, but that they might actually increase adherence. The reasons are not explored in our study, but possibly it is because they better align with physician's clinical experience and international literature and guidelines.

An important strength of our study is that we based the survey on a pilot which consisted of interviews with 10 gastroenterologists, and that the pilot provided insight into which situations led to deviation of the guideline and the reasoning for deviation. However, our study also has three limitations. First, the response rate to the survey was low, which may have led to non-respondent bias. We did not see any differences in age and gender between respondents and the complete group and the number of correctly answered questions did not show a skewed distribution. Still, non-response bias could exist, given that the proportion of responding gastroenterologists was lower in academic hospitals (9.4%) compared to other types of hospital (18.0%). Previous studies have shown that adherence to guidelines is generally higher in academic hospitals implying that we may have underestimated the adherence rate. On the other hand, some respondents requested that an e-learning module of this survey would be developed, indicating that at least some of the responding endoscopists were eager

to improve their knowledge of the guideline and were thus more likely to follow the guideline. In that case, the estimated adherence rate could be overestimated.

Second, we only measured adherence to guidelines among gastroenterologists. This is not a limitation in the Netherlands, because there the vast majority of surveillance endoscopies is performed by gastroenterologists. However, it may hamper the generalizability of our findings to other settings were surveillance endoscopies may also be performed by surgeons or internists. If these clinicians have less knowledge about surveillance guidelines, adherence to guidelines may be lower in these settings.

Finally, our findings are based on a survey, while adherence in daily practice may be different for various reasons. Preferably actual adherence rates are measured. In a survey, gastroenterologists might give desirable answers although they deviate from guidelines in daily practice. Also, if a recommendation is given to a patient, the patient does not always show up after the correct interval.

Our study has four important practical implications. First, the fact that the most often quoted reason for deviation of the guideline is misinterpretation for cases with serrated polyps clearly indicates that the information on these polyps on the score chart or app needs to be improved. Second, it should be further highlighted that according to the guideline HGD should not be taken into account when determining the interval. Moreover, gastroenterologists and pathologists need to discuss how to improve the reporting of the villous or tubulovillous nature of an adenoma in the pathology report to facilitate classification of these lesions. At the time of the introduction of the national colorectal cancer screening programme in 2014, protocols for structured endoscopy and pathology reports were also introduced with predefined categories for histology, which may improve the classification of villous or tubulovillous adenoma. Finally, the use of a pocket-sized score chart, app or other source when making surveillance interval recommendations should be encouraged as this improves adherence to the guideline. Even better would be if, in the future, software could be integrated in the electronic patient dossier and would automatically determine the recommended surveillance interval based on registered polyp characteristics. This would improve the interpretation of the guideline and the compliance to it if it would require a manual override of the system to change this.

The current Dutch guideline differs from other guidelines regarding the level of risk stratification. While other guidelines divide patients in groups based on a simple heuristic using presence or absence of risk factors,[6,21,22] the Dutch guideline combines several risk factors into a score from zero to five. The Dutch guideline is therefore more complex, which may cause misunderstandings and thereby decrease adherence. However, this study shows that more complexity in a guideline did not lower adherence as assessed in a survey, and that this guideline with risk stratification actually

seemed to improve adherence. Since better risk-stratification leads to efficient use of sources and less unnecessary colonoscopies, this should encourage other countries to implement a guideline with more detailed risk-stratification.

In conclusion, the median adherence to the updated colonoscopy surveillance guideline of 76% seems reasonable, and is higher than the adherence to the previous guideline. This shows that detailed (more complex) risk stratification for designation of a surveillance interval is feasible. Adherence could potentially be improved by clarifying the correct interpretation on serrated polyps.

REFERENCES

1. Jemal A, Bray F, Center MM et al. Global cancer statistics. CA Cancer J Clin 2011; 61: 69-90
2. Cottet V, Jooste V, Fournel I et al. Long-term risk of colorectal cancer after adenoma removal: a population-based cohort study. Gut 2012; 61: 1180-1186
3. Leung K, Pinsky P, Laiyemo AO et al. Ongoing colorectal cancer risk despite surveillance colonoscopy: the Polyp Prevention Trial Continued Follow-up Study. Gastrointestinal endoscopy 2010; 71: 111-117
4. Martinez ME, Baron JA, Lieberman DA et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. Gastroenterology 2009; 136: 832-841
5. Atkin WS, Valori R, Kuipers EJ et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. Endoscopy 2012; 44 Suppl 3: Se151-163
6. Lieberman DA, Rex DK, Winawer SJ et al. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2012; 143: 844-857
7. Leddin D, Armstrong D, Borgaonkar M et al. The 2012 SAGE wait times program: Survey of Access to GastroEnterology in Canada. Canadian journal of gastroenterology = Journal canadien de gastroenterologie 2013; 27: 83-89
8. Patel VB, Nahar R, Murray B et al. Exploring implications of Medicaid participation and wait times for colorectal screening on early detection efforts in Connecticut--a secret-shopper survey. Connecticut medicine 2013; 77: 197-203
9. Schreuders EH, Ruco A, Rabeneck L et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015; 64: 1637-1649
10. Terhaar Sive Droste JS, Craanen ME, van der Hulst RW et al. Colonoscopic yield of colorectal neoplasia in daily clinical practice. World journal of gastroenterology 2009; 15: 1085-1092
11. Huppertz J, Coriat R, Leblanc S et al. Application of ANAES guidelines for colonoscopy in France: a practical survey. Gastroenterologie clinique et biologique 2010; 34: 541-548
12. Mulder SA, Ouwendijk RJ, van Leerdam ME et al. A nationwide survey evaluating adherence to guidelines for follow-up after polypectomy or treatment for colorectal cancer. Journal of clinical gastroenterology 2008; 42: 487-492
13. van Kooten H, de Jonge V, Schreuders E et al. Awareness of postpolypectomy surveillance guidelines: a nationwide survey of colonoscopists in Canada. Canadian journal of gastroenterology = Journal canadien de gastroenterologie 2012; 26: 79-84
14. NVMDL. CBO guideline.2014: http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf.
15. Bonnington SN, Rutter MD. Surveillance of colonic polyps: Are we getting it right? World journal of gastroenterology 2016; 22: 1925-1934
16. van Heijningen EM, Lansdorp-Vogelaar I, Kuipers EJ et al. Features of adenoma and colonoscopy associated with recurrent colorectal neoplasia based on a large community-based study. Gastroenterology 2013; 144: 1410-1418
17. de Jonge V, Sint Nicolaas J, van Leerdam ME et al. Systematic literature review and pooled analyses of risk factors for finding adenomas at surveillance colonoscopy. Endoscopy 2011; 43: 560-572
18. Boolchand V, Olds G, Singh J et al. Colorectal screening after polypectomy: a national survey study of primary care physicians. Annals of internal medicine 2006; 145: 654-659

19. Iskandar H, Yan Y, Elwing J et al. Predictors of Poor Adherence of US Gastroenterologists with Colonoscopy Screening and Surveillance Guidelines. Digestive diseases and sciences 2015; 60: 971-978
20. Saini SD, Nayak RS, Kuhn L et al. Why don't gastroenterologists follow colon polyp surveillance guidelines?: results of a national survey. Journal of clinical gastroenterology 2009; 43: 554-558
21. Hassan C, Quintero E, Dumonceau JM et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy 2013; 45: 842-851
22. Matsuda T, Chiu HM, Sano Y et al. Surveillance colonoscopy after endoscopic treatment for colorectal neoplasia: From the standpoint of the Asia-Pacific region. Digestive endoscopy : official journal of the Japan Gastroenterological Endoscopy Society 2016; 28: 342-347

SUPPLEMENTARY MATERIAL

Appendix 1: **Interviews**

10 gastroenterologists were interviewed between May and July 2014. The selected gastroenterologists differed in types and regions of hospitals, gender and age. All interviews were conducted by the same researcher (MvdM) and were audio-recorded. The interviews were semi-structured, started with open questions on which issues were considered advantages and bottlenecks of the guideline by the gastroenterologists. Then, 5 example cases were presented, and recommended surveillance intervals and reasoning were discussed. The response of the interviewees helped to improve the detail of the example cases for the survey and helped to develop several answering options on motivation for deviation from the current guideline.

All gastroenterologists used the new guideline. All gastroenterologists deviated from the recommended interval at least once, one deviated twice, and four deviated three times. Even though all gastroenterologists had a positive opinion about the serrated adenomas being included in the new guideline, 7 gastroenterologists gave the patient with serrated polyps a shorter interval than recommended by the guideline. Although all had a score chart available during the interview, they used it incorrectly, thereby assuming their answer was in accordance with the guideline. We included this in the reasons to deviate from the guideline with the answering option: I was under the assumption that my answer was in line with the guideline.

7 of the 10 gastroenterologists gave a shorter interval for the case with a tubuluvillous adenoma (case 3), 4 did not only score villous, but also tubulovillous adenomas. Six gastroenterologists mentioned that adhering to the guideline was difficult considering the difference between tubulovillous and villous adenomas, because pathology reports in their hospital do not include percentages nor whether adenomas were villous or tubulovillous. All gastroenterologists instructed a 75-year old patient to come back for surveillance. They mentioned that, at the follow-up appointment itself, they would determine whether the patient was healthy enough to undergo colonoscopy. Gastroenterologists mentioned adenoma >20mm is usually not removed by en-bloc polypectomy but it should always be done by piecemeal endomucosal resection (EMR), requiring a surveillance colonoscopy between 4-6 months. This was added to the case (case 3 and 4) accordingly.

3 of the 10 gastroenterologists gave a shorter interval than recommended for the case with high-grade dysplasia (HGD). 8 gastroenterologists mentioned the necessity for clean margins with HGD and 1 gastroenterologist would always recommend surveillance within 3-6 months for a patient with an adenoma with HGD. We therefore added the answering option: based on clinical experience, and based on scientific studies: both

with the sub-answer that there were specific clinical reasons for choosing a different interval.

Appendix 2: **Survey**

Part 2: Cases

Unless noted otherwise:

- All patients were in good health;
- Had no familial risk for colorectal cancer;
- Had undergone their first colonoscopy;
- Bowel preparation was good;
- The cecum was reached;
- The polyp was removed in one piece and endoscopically complete.

	Description case
1	A 60 year old male with 1 distal tubular adenoma of 8mm with low-grade dysplasia
2	A 54 year old male with 1 tubulovillous adenoma of 20mm with low-grade dysplasia in the proximal colon, which was removed by piecemeal. At the subsequent colonoscopy at 6 months no remnant adenomatous tissue was detected.
3	A 69 year old male with 1 distal tubular adenoma of 12 mm with low-grade dysplasia.
4	A 62 year old female with 2 adenomas. Polyp A is a distal tubulovillous adenoma of 10 mm with low-grade dysplasia. Polyp B is a distal villous adenoma of 22 mm with low-grade dysplasia which was removed by piecemeal. At the subsequent colonoscopy at 6 months no remnant adenomatous tissue was detected.
5	A 60 year old female with 5 adenomas. Polyp A is a distal tubular adenoma of 5 mm with low-grade dysplasia. Polyp B is a proximal tubular adenoma of 7 mm with low-grade dysplasia. Polyp C is a proximal tubular adenoma of 4 mm with low-grade dysplasia. Polyp D is a proximal tubular adenoma of 8 mm with low-grade dysplasia. Polyp E is a proximal tubular adenoma of 12 mm with low-grade dysplasia.
6	A 63 year old female with 4 adenomas. Polyp A is a distal tubular adenoma of 6 mm with low-grade dysplasia. Polyp B is a distal tubular adenoma of 5 mm with low-grade dysplasia. Polyp C is a proximal villous adenoma of 9 mm with low-grade dysplasia. Polyp D is a proximal tubular adenoma of 7 mm with low-grade dysplasia.
7	A 79 year old male with 5 adenomas. Polyp A is a distal tubular adenoma of 6 mm with low-grade dysplasia. Polyp B is a proximal tubular adenoma of 8 mm with low-grade dysplasia. Polyp C is a proximal tubular adenoma of 6 mm with low-grade dysplasia. Polyp D is a proximal tubular adenoma of 4 mm with low-grade dysplasia. Polyp E is a distal tubular adenoma of 5 mm with low-grade dysplasia.
8	A 75 year old male with 4 adenoma. Polyp A is a distal tubular adenoma of 4 mm with low-grade dysplasia. Polyp B is a distal tubular adenoma of 6 mm with low-grade dysplasia. Polyp C is a distal tubular adenoma of 12 mm with high-grade dysplasia. Polyp D is a distal tubular adenoma of 9 mm with low-grade dysplasia.
9	A 65 year old male with 1 distal tubulovillous adenoma of 11 mm with high-grade dysplasia.
10	A 58 year old female with 1 serrated adenoma of 8 mm in the proximal colon.
11	A 54 year old female with 2 polyps. Polyp A is a proximal sessile serrated adenoma/polyp of 10 mm. Polyp B is a proximal sessile serrated adenoma/polyp of 12 mm.

	Description case
12	A 51 year old male with 2 adenomas. Polyp A is a distal tubular adenoma of 7 mm with low-grade dysplasia. Polyp B is a distal villous adenoma with low-grade dysplasia. The male has a brother who was diagnosed with colorectal cancer at age 48 and in which no hereditary syndrome was diagnosed by the clinical geneticist.
13	A 53 year old male with 1 distal tubulovillous adenoma of 11 mm with low-grade dysplasia. His sister was diagnosed with colorectal cancer at age 45 years old.
14	A 69 year old male with a negative colonoscopy. However, the male has had one previous colonoscopy where 2 adenomas were detected. Polyp A was a distal villous adenoma of 12 mm with low-grade dysplasia. Polyp B was a proximal tubular adenoma of 8 mm with low-grade dysplasia.
15	A 63 year old female with a negative colonoscopy. Five years ago she also had a negative colonoscopy. She underwent this colonoscopy because she had 2 detected adenomas at a previous colonoscopy. Polyp A was a distal tubular adenoma of 5 mm with low-grade dysplasia. Polyp B was a distal tubular adenoma of 12 mm with low-grade dysplasia.

What surveillance recommendation would you provide?

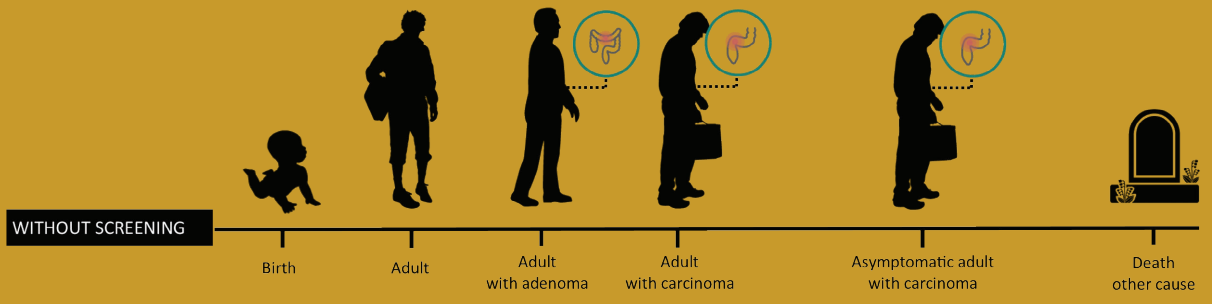
- interval of <1 year
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
- 10 years
- no surveillance
- no surveillance, unless in good condition, then in 3 years
- no surveillance, unless in good condition, then in 5 years
- no recommendation yet, but referral to the clinical geneticist.

Part 3: reasons for deviation

Why did you deviate from the guideline in this case?

- I thought that my answer was in agreement with the guideline;
- I did not read the question correctly;
- I am not familiar with the new guideline;
- I based my recommendation on scientific evidence;
- I based my recommendation on clinical experience;
- Other (+ free text field).


WITH SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Screening
Patient
Death other cause


WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Asymptomatic adult with carcinoma
Death other cause

Chapter 8

General discussion

In this discussion, the research questions of Chapter 1 will be answered. Subsequently, a few methodological aspects will be addressed. Finally, the implications will be discussed and recommendations for future research are given.

ANSWERS TO THE RESEARCH QUESTIONS

Are there differences in FIT performance between men and women and do men and women need to be screened differently with fecal immunochemical testing from a cost-effectiveness perspective?

Men and women do not have significantly different positive predictive values for most cut-off levels, resulting in a similar harm-to-benefit ratio. Men do, however, have higher positivity rates than women, reflected by both higher detection rates and a higher false positive rate. A higher false-positive rate in men implies that specificity is lower in men than in women. Despite these differences in performance, our model showed that screening stratified by gender does not improve cost-effectiveness. Therefore, our findings support uniform screening of men and women as currently applied in most FIT screening programs, like in the Netherlands.

Several countries around the world have adopted the faecal immunochemical test for haemoglobin (FIT) for population-based screening. An important topic of debate has been whether these programmes should be tailored by gender. In Chapter 2, data from a Dutch pilot study were used to show FIT performance, in terms of yield and positive predictive value, per gender. In Chapter 3, these values were used in the MISCAN-Colon model to estimate gender specific FIT sensitivity and specificity. The estimated sensitivity in women was lower than in men, while the estimated specificity was higher in women. We showed that even if sensitivity of FIT is lower in women than in men and FIT screening therefore yields less benefit compared to no screening, the incremental costs and benefits of more intensive screening compared to less intensive screening is similar for both genders. Moreover, the majority of the optimal FIT screening strategies are identical for men and women. Consequently, the QALYs gained for uniform screening and screening stratified by gender never differ more than 7% and are equal for willingness-to-pay thresholds of more than €1300. In other words, there is little to no benefit from FIT screening stratified by gender compared to uniform screening, and our findings support the current policy of uniform FIT screening.

Do systematic false-negative fecal immunochemical test results exist and what are their implications for screening effectiveness?

A percentage of adenomas are systematically missed by repeated FIT screening, presumably due to nonbleeding adenomas. This phenomenon lowers the impact of

FIT screening on mortality reduction by an estimated 6.4%. In addition, a proportion of individuals will systematically test false-positive on FIT screening. Since these individuals will not receive screening for a period after their negative colonoscopy, this phenomenon lowers the number of (expected) unnecessary colonoscopies.

Long-term effectiveness of FIT screening has been estimated using modeling, but these models all assumed that FIT results are independent of each other over screening rounds. However, if some adenomas do not bleed over several years, they will cause systematic false-negative FIT results. Observed data from two FIT screening rounds of the CORERO study enabled us to further evaluate this issue. In Chapter 4, we compared observed second round adenoma detection rates with rates simulated by the MISCAN-Colon model and estimated which proportion of adenomas is systematically missed to explain the observed rates. Subsequently, we estimated the effect this would have on the effectiveness of a FIT screening program. Chapter 4 shows that over 70% of non-advanced adenomas and 26% of advanced adenomas have to be systematically missed to explain the observed rates. We estimated that this will impair the effectiveness of a FIT screening program with approximately 6%.

What is the comparative cost-effectiveness of CTC versus colonoscopy screening with assumed data on attendance and costs from a randomized controlled screening trial in a dedicated screening setting?

Based on the 56% higher attendance of CTC compared to colonoscopy as observed in a randomized controlled trial, CTC screening for colorectal cancer is more cost-effective than colonoscopy screening. If an individual will participate independent of the screening test, it is most cost-effective to screen with colonoscopy.

Earlier cost-effectiveness analyses comparing colonoscopy with CT-colonography (CTC) screening showed that the results were highly sensitive to assumptions regarding test costs and attendance. At that time comparative data for these assumptions were lacking. In Chapter 5, we used data from the COCOS-trial comparing CTC and colonoscopy in a dedicated screening setting to perform a representative cost-effectiveness analysis. We used the microsimulation model MISCAN-Colon to compare the Quality Adjusted Life years (QALYs) gained and the costs of several screening strategies with CTC and colonoscopy. For both tests we determined optimal age range and interval combinations with 100% attendance. For these combinations, we compared the cost-effectiveness of both tests assuming observed attendance. Chapter 5 shows that because of the higher attendance rates, CTC screening is more cost-effective than colonoscopy screening. However, the implementation of CTC screening requires prior satisfactory resolvement of how to deal with extracolonic findings.

What are the socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal screening programme with faecal immunochemical testing?

Faecal immunochemical testing resulted in higher detection of advanced neoplasia in more deprived participants. Therefore, it has the potential to reduce existing socioeconomic inequalities in CRC mortality. However, this higher detection is currently offset by the lower participation in this group. Further research is needed into this lower participation to ensure well-informed decision-making.

CRC mortality rates are higher for individuals with a lower socioeconomic status (SES). Screening could influence health inequalities. In Chapter 6, we therefore investigated SES differences in participation and diagnostic yield of the Dutch FIT screening programme in 2014 and 2015. 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. We showed lower participation to FIT for Quintile 5 (67.0%) compared to the other Quintiles (73.0% to 75.1%), while the detection rate per FIT participant for advanced neoplasia gradually increased from 3.3% in Quintile 1 to 4.0% in Quintile. 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%). Therefore, the yield per invitee does not significantly differ for SES.

How do gastroenterologists interpret and comply to the risk-stratified guideline for surveillance after polypectomy?

The median compliance to the risk-stratified colonoscopy surveillance guideline of 76% seems reasonable, and is higher than the compliance to the previous guideline. This shows that detailed (more complex) risk stratification for designation of a surveillance interval is feasible. Compliance could potentially be improved by clarifying the correct interpretation on serrated polyps.

Low compliance to the Dutch guideline for colonoscopy surveillance after polypectomy led to the release of a new guideline in 2013. This new guideline was risk-stratified at a more detailed level than the previous one to achieve more efficient use of colonoscopy resources. In Chapter 7, we assessed the feasibility of the risk-stratified guideline by evaluating the correct interpretation of and compliance to this guideline. Based on semi-structured interviews with 10 gastroenterologists, we developed an online survey to evaluate gastroenterologists' recommendations for surveillance in 15 example cases of patients with polyps. If their recommended intervals deviated from the new guideline, respondents were asked to indicate their motives for doing so. Gastroenterologist gave a correct recommendation in a median of 10 out of 15

cases. The percentage of correct recommendations per case ranged from 14% to 95% (median case 76%). Deviations were mainly due to misinterpretation of the guideline with respect to serrated polyps (48%) or misreading of the questions (30%). Chapter 7 shows that detailed (more complex) risk stratification for designation of a surveillance interval is feasible. Compliance could potentially be further improved by clarifying the correct interpretation on specific aspects of the guideline.

METHODOLOGICAL ISSUES

Some methodological issues in the papers could have influenced the results of the papers. In this paragraph, I describe three of the most important methodological issues of this thesis, how the issues were handled and why and how they could have influenced the results. Two of these issues are assumptions in the MISCAN-model (the size distribution of non-progressive adenomas and the attendance (over time) to screening tests). The other issues concern participation bias: in the survey on the adherence to the guideline, in the national screening program and as modeled in MISCAN-Colon.

Size distribution of adenomas

Until the analysis of Chapter 5, the size distribution of adenomas in the MISCAN-Colon model had been based on autopsy studies. [1-11] When we validated the MISCAN-Colon model to observed detection rates of colonoscopy in the Dutch COCOS trial, we noticed that the overall prevalence of adenomas of the model was in concordance with the observed prevalence, but the size distribution of adenomas was markedly different. The COCOS trial showed a substantially higher percentage of small adenomas (≤5mm) than the model (69.5% versus 47.6%), and a lower percentage of medium (6-9mm) (15.9% versus 34.9%) and large adenomas (≥10mm) (14.6% versus 17.7%) (Table 1 and 2).

However, the size distribution of adenomas detected in COCOS did not show any differences with those detected in other colonoscopy studies (Table 2),[12-15] suggesting that the size of adenomas is different when measured during colonoscopy than during autopsy.

Different arguments can be used to determine which size distribution is the best to use: for instance it is harder to measure the adenoma during colonoscopy, since the scope might not always measure the lesion at a 90 degrees angle, resulting in an underestimation of the size of the adenoma. However, obviously, the adenoma is in situ during coloscopy, while the removal of the colon at autopsy might influence the size of the tissue. In addition, the colonoscopy studies are far larger and more recent than the autopsy studies.

Another argument to choose the size distribution based on the COCOS study is that the calibrated FIT sensitivity turned out to be more in line with published data on FIT sensitivity of advanced adenomas (Table 3)[16-20] than the calibrated FIT sensitivity when assuming the size distribution based on autopsy studies. The sensitivity of FIT in MISCAN-Colon is calibrated to observed detection rates and will be higher if fewer adenomas are present in our model. Since fewer large and medium adenomas are present in MISCAN-Colon with the new size distribution, the calibrated FIT sensitivity for advanced neoplasia increased to be more in line with the other published studies.

Therefore, we chose to assume the size distribution of the colonoscopy trial COCOS as the base case analysis in all Chapters. The alternative size distribution based on autopsy studies was used in sensitivity analyses.

The size distribution of adenomas was adjusted only in the growth of nonprogressive adenomas, since the growth of progressive adenomas is calibrated to the incidence and mortality reduction of another endoscopy study, the Atkin trial.[21] Since the number and growth of progressive adenomas did not change, the adjustment in size distribution of adenomas will have little effect on the (cost)effectiveness of endoscopy and CTC screening. However, the (cost)effectiveness of FIT screening in our studies is increased by the increase in sensitivity in medium and large adenomas, and therefore the comparative effectiveness of FIT to endoscopy and CTC screening is also increased by the change in size distribution of adenomas.

Furthermore, the new size distribution influenced the results of the proportion of adenomas that are systematically missed by repeated FIT screening (Chapter 4). The proportion of adenomas that are systematically missed by repeated FIT screening was calibrated to the observed detection rate of the second round. With the new size distribution, fewer large and medium adenomas are present in MISCAN, while the assumed FIT sensitivity increased. Therefore, the modeled second round detection rate will be lower and a lower proportion of adenomas that are systematically missed was necessary after implementing the new size distribution.

Table 1 *The proportions of small, of medium adenomas and of large adenomas in all of the detected adenomas across different autopsy studies, in the COCOS colonoscopy arm and as simulated in MISCAN-Colon*

Study	% small adenomas	% medium adenomas	% large adenomas
COCOS 1*	69.5%	15.9%	14.6%
COCOS 2*	**73.1%**	**14.7%**	**12.1%**
Rickert (table 4)	43.7%	40.9%	15.4%
Williams (table 5)	42.6%	44.6%	12.8%
Vatn (table 4)	31.8%	53.7%	14.4%
Jass (table 8)	38.9%	43.0%	18.1%
Bombi (table 3)	46.1%	47.2%	6.7%
Arminski (table 6)	58.6%	26.1%	15.3%
Blatt (table 4)	49.9%	34.0%	16.1%
Total autopsy studies	47.6%	34.9%	17.5%
MISCAN-Colon 2a**	**47.6%**	**34.7%**	**17.7%**
MISCAN-Colon 2b**	**73.1%**	**15.0%**	**11.9%**

Small adenomas are defined as adenomas < 5mm, medium adenomas as 5-9 mm and large adenomas as ≥10 mm.
* COCOS as observed (1) and corrected for sensitivity of colonoscopy (2)
** MISCAN-Colon size distribution of adenomas based on autopsy studies (2a), MISCAN-Colon based on COCOS (2b).

Table 2 *The proportion of individuals with a small or medium adenoma and of individuals with at least one large adenoma across different colonoscopy studies and as simulated in MISCAN-Colon.*

Study	Age range	% male	N	% small/ medium	% large
COCOS	**50-75**	**0.5**	**1276**	**78.8%**	**21.2%**
Regula	40-49	0.359	7106	78.8%	21.2%
Regula	50-66	0.359	43024	76.3%	23.7%
Lieberman	mean 62.9	0.968	3121	76.9%	23.1%
Barclay	60.3	0.51	2053	78.2%	21.8%
Chen, Rex	56	0.47	10034	84.8%	15.2%
MISCAN-Colon 2a*	**50-75**	**na**	**na**	**66.6%**	**33.4%**
MISCAN-Colon 2b*	**50-75**	**na**	**na**	**76.4%**	**23.6%**

adenomas are defined as adenomas < 5mm, medium adenomas as 5-9 mm and large adenomas as ≥10 mm.
* MISCAN-Colon 2a size distribution of adenomas based on autopsy studies, MISCAN-Colon 2b based on COCOS.

Table 3. *Sensitivity for FIT (OC-sensor) with a cut off of 10 ug Hb/g feces in two versions of MISCAN and in several published studies.*

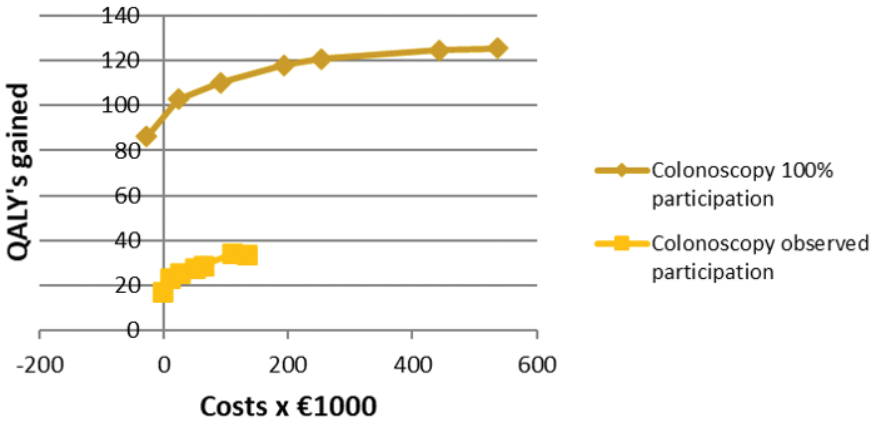
Source	Sensitivity for advanced adenomas (%)
MISCAN-Colon 2a	**16.2**
MISCAN-Colon 2b	**34.4**
Imperiale (meta-analysis)	40 (33-47)
van Wijkerslooth	35(26–45)
Kim	44(39–49)
Hernandez	32(22–42)
Chang	32(27–37)
Khalid-de Bakker	16(6–31)

** MISCAN-Colon size distribution of adenomas based on autopsy studies (2a), MISCAN-Colon 2 based on COCOS (2b).

Participation (over time)

The assumed participation in a model has a large influence on the estimated costs and effects of a screening program. In Chapter 3 and 5, we determined the most optimal combinations of age range and interval of a screening test assuming 100% participation. However, to compare the screening tests with each other, the most plausible situation should be modeled. When screening strategies were considered with realistic participation, both costs and QALYs gained decreased compared to full attendance (see an example of various colonoscopy strategies assuming 100% and realistic attendance in Figure 1). This decrease in costs and QALYs gained is approximately proportionally to the decrease in attendance.

Figure 1. *The costs and QALY gained of the most optimal colonoscopy strategies assuming 100% participation and realistic participation (21.5%).*



In this thesis, the assumed participation rates of the different screening tests are all based on Dutch pilot studies with a similar set-up and could therefore be reliably compared to each other. However, test preferences and resulting participation could change over time, which could significantly influence the comparative cost-effectiveness of different screening tests. Therefore, it is important to continue to study patient preferences to the different tests.

Not only the overall attendance has a clear influence on the outcomes of screening, also the pattern of attendance over several screening rounds can change the effectiveness of a screening program. If the participation of individuals over several screening rounds is random, the result will be that (almost) all individuals receive some screening. In contrast, if participation is clustered in subpopulations, the same individuals show up for screening every time. The result is that these individuals receive regular screening, while others receive no screening at all. Population screening is more effective if participation over rounds is random, because first (prevalent) screens have a higher yield than subsequent (incident) screens.

In Chapter 5 we added a sensitivity analysis in which we assume two extreme scenarios for subsequent participation: that the participation over several rounds is equally distributed (random next participation) or fixed in a subpopulation (fixed next participation). For example, assuming random participation of 50%, an individual has a 50% probability to participate each screening round. Thus, individuals invited to 5 screening rounds only have a 3% (0.5^5) probability of skipping all screening invitations, meaning that 97% of invited individuals will participate in at least one screening round. However, assuming fixed attendance of 50%, the same 50% of invited individuals attend all screens, while the other 50% is not screened at all.

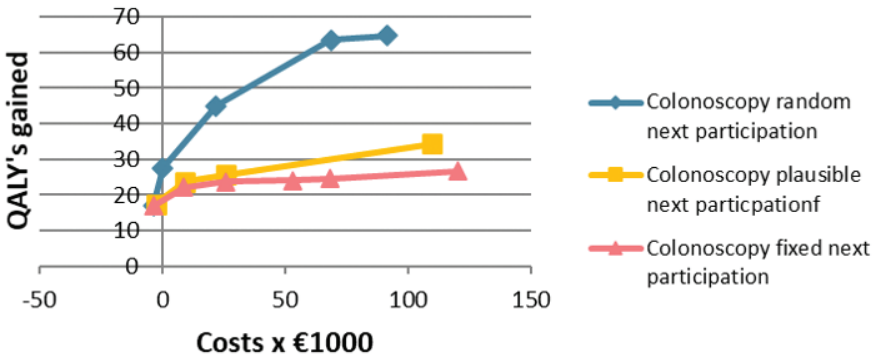
Figure 2 shows the large influence of these alternative scenarios on the costs and effects of colorectal cancer screening, and also shows the influence on the comparative effectiveness of colonoscopy versus CTC. The reality will lie somewhere between these extreme scenarios. The motive of an individual not to participate in initial screening could also be a motive not to participate in subsequent rounds, for instance because they worry about unpleasantness of the test or risks.[22] Therefore, individuals that did not attend a screening are likely to have a smaller probability to attend the next screening round. Vice versa, it is also likely that individuals who participate also have a higher probability to participate in subsequent screening round(s).

In this thesis, we therefore assumed a higher probability to attend screening for individuals that previously attended screening, based on published FIT participation over two screening rounds. This scenario approaches the fixed next attendance; but the non-attenders do have a probability to attend the next screening round. Three

subgroups are now used in the MISCAN-Colon model, previous attenders, previous non-attenders and never-attenders. In reality, however, more subgroups probably exist.

Because of the influence on costs and effects, it will be interesting to gather more data on the pattern of participation in individuals over several round, for instance in the current Dutch national screening program.

Figure 2. *The costs and QALY gained of the most optimal colonoscopy strategies with random, plausible and fixed next participation.*



Participation bias

In addition to how many individuals attend screening, the characteristics of the individuals that participate in screening will have an influence on the effectiveness of a screening program. If participating individuals are not representative for the invited population because they have certain traits that could affect the outcome, participation bias or non-response bias could occur. In primary and secondary prevention, participation bias often exists, primarily because the people with a higher risk of the disease have a lower participation rate than people with a lower risk of the disease. Participation bias can influence results of screening studies in several ways. For instance, if participants have a lower risk of colorectal cancer, a lower colorectal cancer rate in these individuals compared to non-participants is (partly) due to their lower risk and not (totally) attributable to screening.

In MISCAN-Colon participation bias could also influence model outcomes. If the (cost)effectiveness of a total screening program is modeled assuming participants have an equal risk at CRC as the total population, the program effectiveness might be overestimated. In contrast, the sensitivity of FIT could be underestimated if the screened population has a lower risk of colorectal cancer (and therefore also adenomas) than the total population. As described earlier in the discussion, FIT sensitivity is calibrated to observed detection rates, and a valid calibration warrants a good estimate of the

prevalence pool of adenomas in the screened population. To account for possible participation bias in analyses performed with MISCAN-Colon, the incidence of CRC of possible participants is assumed to be lower than the incidence in the total population, based on a randomized controlled trial with gFOBT screening.[23]

One characteristic of screening participants that has been shown to be different is socioeconomic status. In Chapter 6, we showed a difference in participation by socioeconomic status in the Dutch colorectal cancer screening program. In addition to the main conclusions of the chapter, indications of participation bias can be drawn from the chapter because we analysed the background incidence of the population. Interestingly, no socioeconomic differences in background incidence were found. This could indicate that socioeconomic status is not a trait resulting in participation bias. The detection rate of FIT did differ for SES. This can be caused by a higher FIT sensitivity in lower SES, which seems unlikely because a previous study observed a similar stage distribution of screen-detected CRC across SES quintiles.[24] A more likely explanation is that participation bias exists within the SES quintiles, for instance, 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. Differences in participation between native Dutch and ethnic minorities have been previously reported.[25]

Participation bias can occur in all types of studies and in Chapter 7, the low response of gastroenterologists to the survey could impose a risk for non-response (participation) bias.

To assess this, we compared the characteristics of respondents and non-respondents and we did not see any differences in age group and gender. We did observe a difference in response between gastroenterologists in an academic hospital (9.4%) or a different type of hospital (18.0%). Previous studies have shown that adherence to guidelines is generally higher in academic hospitals, implying that we may have underestimated the compliance to the guideline. In addition, gastroenterologists with a strong opinion (either positive or negative) might have been more prone to participate, irrespective of age, gender and type of hospital. This could have been observed in the data if a subgroup of gastroenterologists had a very low compliance while at the same time another subgroup of gastroenterologists had a very high compliance. This would result in a skewed distribution in the number of correctly answered questions, while the distribution in the study was closer to a normal distribution. Some requests were received to develop an e-learning module of this survey, indicating that at least some of the responding endoscopists were eager to improve their knowledge of the guideline and were thus more likely to follow the guideline. The estimated adherence rate to the

guidelines should then be considered as an upper bound for the adherence rate among gastroenterologists in the Netherlands at large.

IMPLICATIONS

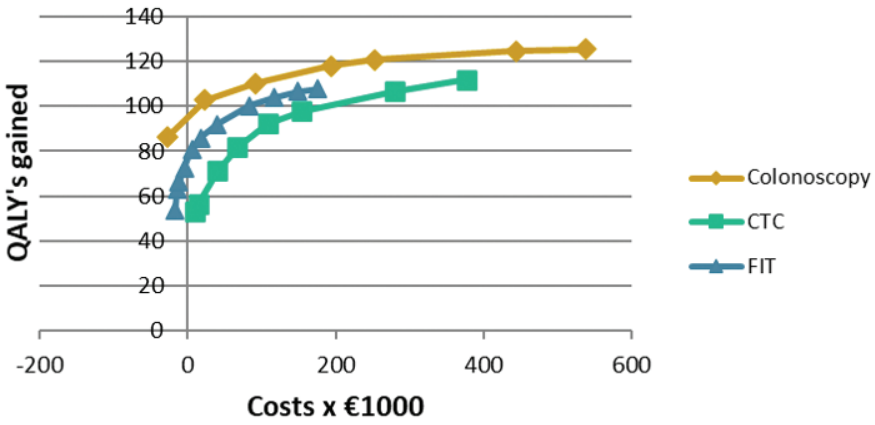
Costs-effectiveness of screening tests

The cost-effectiveness of three different screening tests was estimated in this thesis, but they were not all compared with each other. The results of Chapter 3 and Chapter 5 combined show the comparison of the cost-effectiveness of CTC, colonoscopy and FIT (Figure 3 and Figure 4). There were some slight differences in model assumptions between the chapters: in Chapter 5, we assumed somewhat lower utility losses due to colonoscopy compared to Chapter 3 and the analysis of Chapter 3 is performed with a separate male and a female model, while the analysis of Chapter 5 is performed with a model of the total population. However, these are not expected to influence the comparison in a substantial way.

Figure 3 shows that for participating individuals, the cost-effectiveness of colonoscopy and FIT is comparable and both tests dominate CTC screening. Systematic FIT results have a negative effect on FIT effectiveness, which causes colonoscopy screening to be slightly more cost-effective. With the participation as observed in the Dutch pilots (60% to FIT, 34% to CTC and 22% to colonoscopy), FIT screening is most cost-effective. In figure 4, the different FIT scenarios with or without systematic FIT results and gender-specific screening are also displayed, showing that the impact of these factors on the FIT effectiveness is small compared to the differences in comparative effectiveness with CTC and colonoscopy.

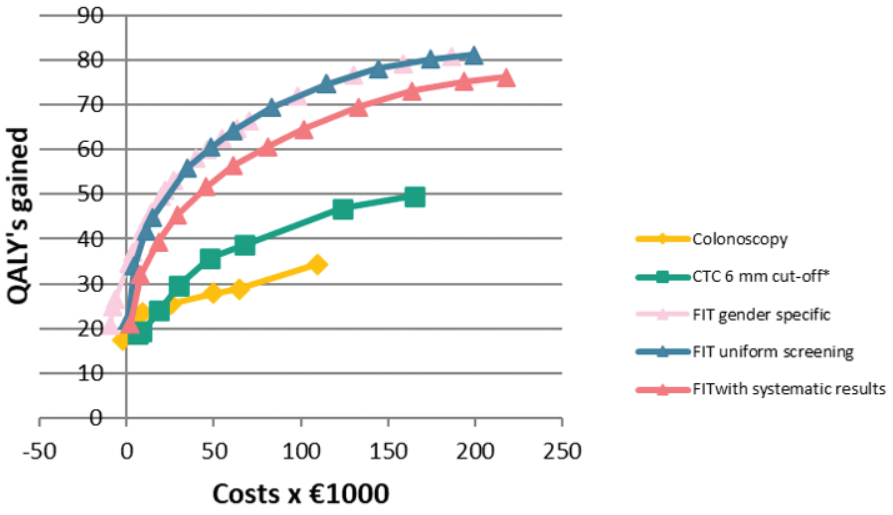
Another option as a screening test for CRC screening is sigmoidoscopy. For participants this is shown to have a comparable mortality reduction as both annual FIT screening and 1-time colonoscopy.[26] The participation to sigmoidoscopy has been shown to be inferior to FIT and comparable to CTC screening.[27, 28] With this participation, it is highly likely that sigmoidoscopy will not be a cost-effective alternative compared to FIT.

Figure 3. *The costs and QALY gained of the most optimal colonoscopy, CTC and FIT strategies assuming perfect (100%) participation.*



Other screening tests for CRC screening are also under development, for instance biomarker tests. These biomarkers tests could increase the accuracy to detect preclinical adenomas and CRC, while at the same time maintaining a high participation rate because the method of sample collection is very similar to FIT. A study (with MISCAN-Colon) showed that, assuming 53% sensitivity for adenomas and 100% sensitivity for CRC and 100% specificity, costs of such a test should not exceed 7 times the costs of FIT in order to be cost-effective.[29] The test costs could be higher if participation of such a test would exceed FIT participation, if adenoma sensitivity would be higher than assumed or if colonoscopy capacity is limited (because of a higher PPV).

Figure 4. *The costs and QALY gained of the most optimal colonoscopy, CTC and FIT strategies assuming realistic participation.*



Surveillance

Even though the content of the guideline has been updated to accommodate scientific expertise, the impact of these changes on long-term (costs)effectiveness is limited if both guidelines would have perfect adherence to the guideline, as shown in Table 4. In contrast, the adherence to the guideline itself has a far larger impact on the cost-effectiveness. Therefore, it is very important that adherence to the new guideline for surveillance after polypectomy is as high as or higher than adherence to the older version. Chapter 7 showed that, even though the guideline has (more complex) risk stratification, the adherence seemed to improve compared to the old guideline if tested in a survey. It also provided direct practical implications to further increase the gastroenterologist's adherence to the guideline: the information on how to score serrated polyps and HGD should be further clarified and the use of a pocket-sized score chart, app or other source when making surveillance interval recommendations should be encouraged.

To ensure that gastroenterologists will also adhere to the guideline in practice, it could be even better if in the future, software could be integrated in the electronic patient records and would automatically determine the recommended surveillance interval based on registered polyp characteristics. This would improve the interpretation of the guideline and the compliance to it if it would require a manual override of the system to change the recommended interval. If this system could also send automatic

reminders to a planner or the patient around the recommended interval, it might also increase patient's adherence.

Adherence to the guideline is tested in a survey in Chapter 7. In practice, the actual adherence rates could differ for various reasons. For instance, the gastroenterologists' decisions can differ in practice if they gave a desirable answer in the survey or if other factors than included in the survey impact the gastroenterologists' decision. Also, if a recommendation is given to a patient, the patient does not always show up after the correct interval. Therefore, it is recommended to estimate actual adherence rates in future research and compare it to the adherence to the old guideline.

Table 4. *Surveillance outcome (undiscounted) assuming adherence as in practice, assuming perfect adherence to the old guideline and perfect adherence to the new guideline.*

Scenario	Prevented CRC cases	Prevented CRC deaths	QALY gained *	Netto costs (€) *	Number of colonoscopies	Number needed to scope to prevent 1 death	Life days gained per colonoscopy
Old practice	52	31	135	633,909	5574	177	21
Old guideline assuming perfect adherence	51	34	147	-97,761	2336	69	51
New guideline assuming perfect adherence	47	33	144	-97,381	1904	57	60

Personalised or risk-stratified screening

Chapter 3 showed that screening stratified by gender is not likely to be cost-effective. However, other personal factors could have a larger impact on the (cost)effectiveness of screening, such as race, comorbidity, exposure to risk factors for colorectal cancer and previous screening history (FIT results and non-participation). A previous article with MISCAN-Colon showed that comorbidity, for example, could halve the benefits from screening if all other factors of the individual are equal.[30]

Individualized screening could further optimize screening by improving screening effectiveness while reducing the probability of harm and resource utilization. However, this can only happen if participation to (the individual) screening program is maintained. In Chapter 7, we showed that for a gastroenterologist a more detailed risk stratification in surveillance is feasible, indicating it is also feasible to perform in screening. Patients preferences for individualized screening will be important, for instance what will happen if a person is offered less intensive screening than their peers. Future research to patient preferences in individualized or risk-stratified screening and the impact of

certain individual characteristics on the (cost)effectiveness of screening is warranted. In addition, cost-effectiveness analysis as performed in Chapter 3 can help to identify if personalised screening based on other individual factors is cost-effective compared to uniform screening.

CONCLUSIONS

- Men have higher positivity rates than women, reflected by both higher detection rates and a higher false positive rate. (Chapter 2)
- Despite sex differences in FIT performance and CRC risk, CRC screening stratified by sex does not improve cost-effectiveness. (Chapter 3)
- A comparison between observed and modeled second round detection rates shows that a proportion of adenomas is systematically missed by repeated FIT screening. This impairs the efficacy of FIT screening slightly. (Chapter 4)
- Based on randomized controlled screening trial data, showing higher participation rates for CTC, CTC is more cost-effective than colonoscopy screening for colorectal cancer. (Chapter 5)
- 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 program, this is currently offset by the lower participation in this group. (Chapter 6)
- The adherence to the updated risk-stratified colonoscopy surveillance guideline seems reasonable, and is higher than the adherence to the previous simple guideline. A detailed (more complex) risk stratification for designation of a surveillance interval is therefore feasible. (Chapter 7)

RECOMMENDATIONS

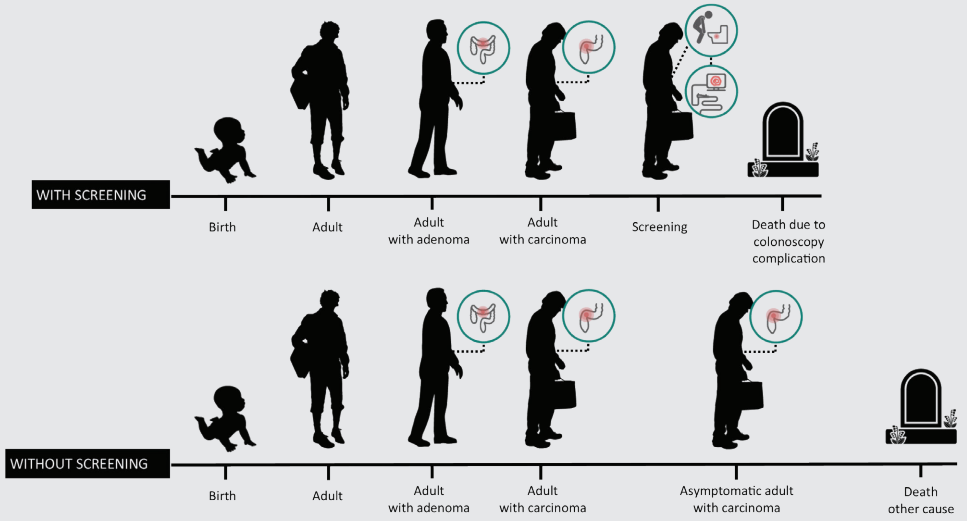
- In case of sufficient colonoscopy capacity, FIT screening should not be stratified by sex. (Chapter 2 and 3)
- The program performance of FIT screening over multiple screening rounds should be estimated in future studies to validate the dependency of FIT test results. (Chapter 4)
- CTC should be included in future studies that evaluate alternative screening tests for the FIT. (Chapter 5)
- Further research is needed to clarify on how to target socioeconomic groups that are less compliant and/or more at risk for CRC and ensure well-informed decision-making. (Chapter 6)

- To improve adherence to the guideline for surveillance after polypectomy, clarification on how to score serrated adenomas and high-grade dysplasia should be given, and the use of a pocket card and/or app should be further encouraged. (Chapter 7)

REFERENCES

1. Arminski TC, McLean DW. Incidence and Distribution of Adenomatous Polyps of the Colon and Rectum Based on 1,000 Autopsy Examinations. Dis Colon Rectum. 1964;7:249-61.
2. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. Cancer. 1988;61(7):1472-6.
3. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. Ann Surg. 1963;157:223-6.
4. Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, Jensen OM, Koskela E, MacLennan R, Simpson JG, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. Int J Cancer. 1985;36(2):179-86.
5. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. Scand J Gastroenterol. 1989;24(7):799-806.
6. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet (London, England). 1996;348(9040):1467-71.
7. Morson B. President's address. The polyp-cancer sequence in the large bowel. Proc R Soc Med. 1974;67(6 Pt 1):451-7.
8. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer. 1975;36(6):2251-70.
9. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. Cancer. 1979;43(5):1847-57.
10. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. Cancer. 1982;49(4):819-25.
11. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. Gut. 1982;23(10):835-42.
12. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. The New England journal of medicine. 2006;355(24):2533-41.
13. Chen SC, Rex DK. Variable detection of nonadenomatous polyps by individual endoscopists at colonoscopy and correlation with adenoma detection. Journal of clinical gastroenterology. 2008;42(6):704-7.
14. Lieberman D, Moravec M, Holub J, Michaels L, Eisen G. Polyp size and advanced histology in patients undergoing colonoscopy screening: implications for CT colonography. Gastroenterology. 2008;135(4):1100-5.
15. Regula J, Rupinski M, Kraszewska E, Polkowski M, Pachlewski J, Orlowska J, Nowacki MP, Butruk E. Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia. The New England journal of medicine. 2006;355(18):1863-72.
16. de Wijkerslooth TR, Stoop EM, Bossuyt PM, Meijer GA, van Ballegooijen M, van Roon AH, Stegeman I, Kraaijenhagen RA, Fockens P, van Leerdam ME, Dekker E, Kuipers EJ. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia. The American journal of gastroenterology. 2012;107(10):1570-8.
17. Hernandez V, Cubiella J, Gonzalez-Mao MC, Iglesias F, Rivera C, Iglesias MB, Cid L, Castro I, de Castro L, Vega P, Hermo JA, Macenlle R, Martinez-Turnes A, Martinez-Ares D, Estevez P, Cid E, Vidal MC, Lopez-Martinez A, Hijona E, Herreros-Villanueva M, Bujanda L, Rodriguez-Prada JI. Fecal immunochemical test accuracy in average-risk colorectal cancer screening. World journal of gastroenterology. 2014;20(4):1038-47.

18. Imperiale TF, Gruber RN, Stump TE, Emmett TW, Monahan PO. Performance Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis. Annals of internal medicine. 2019;170(5):319-29.
19. Khalid-de Bakker CA, Jonkers DM, Sanduleanu S, de Bruine AP, Meijer GA, Janssen JB, van Engeland M, Stockbrugger RW, Masclee AA. Test performance of immunologic fecal occult blood testing and sigmoidoscopy compared with primary colonoscopy screening for colorectal advanced adenomas. Cancer prevention research (Philadelphia, Pa). 2011;4(10):1563-71.
20. Kim NH, Park JH, Park DI, Sohn CI, Choi K, Jung YS. The fecal immunochemical test has high accuracy for detecting advanced colorectal neoplasia before age 50. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2017;49(5):557-61.
21. Atkin WS, Valori R, Kuipers EJ, Hoff G, Senore C, Segnan N, Jover R, Schmiegel W, Lambert R, Pox C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. Endoscopy. 2012;44 Suppl 3:Se151-63.
22. van Dam L, Korfage IJ, Kuipers EJ, Hol L, van Roon AH, Reijerink JC, van Ballegooijen M, van Leerdam ME. What influences the decision to participate in colorectal cancer screening with faecal occult blood testing and sigmoidoscopy? European journal of cancer (Oxford, England : 1990). 2013;49(10):2321-30.
23. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. Cancer. 1986;58(2):397-403.
24. Toes-Zoutendijk E, Kooyker AI, Elferink MA, Spaander MCW, Dekker E, Koning HJ, Lemmens VE, van Leerdam ME, Lansdorp-Vogelaar I. Stage distribution of screen-detected colorectal cancers in the Netherlands. Gut. 2017.
25. de Klerk CM, Gupta S, Dekker E, Essink-Bot ML. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. Gut. 2017.
26. Buskermolen M, Cenin DR, Helsingen LM, Guyatt G, Vandvik PO, Haug U, Bretthauer M, Lansdorp-Vogelaar I. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a microsimulation modelling study. BMJ (Clinical research ed). 2019;367:l5383.
27. Hol L, Kuipers EJ, van Ballegooijen M, van Vuuren AJ, Reijerink JC, Habbema DJ, van Leerdam ME. Uptake of faecal immunochemical test screening among nonparticipants in a flexible sigmoidoscopy screening programme. Int J Cancer. 2012;130(9):2096-102.
28. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, van de Vijver MJ, Biermann K, Thomeer M, van Leerdam ME, Fockens P, Stoker J, Kuipers EJ, Dekker E. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. The Lancet Oncology. 2012;13(1):55-64.
29. Lansdorp-Vogelaar I, Goede SL, Bosch LJW, Melotte V, Carvalho B, van Engeland M, Meijer GA, de Koning HJ, van Ballegooijen M. Cost-effectiveness of High-performance Biomarker Tests vs Fecal Immunochemical Test for Noninvasive Colorectal Cancer Screening. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2018;16(4):504-12.e11.
30. Saini SD, van Hees F, Vijan S. Smarter screening for cancer: possibilities and challenges of personalization. Jama. 2014;312(21):2211-2.


WITH SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Screening
Death due to colonoscopy complication
WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Asymptomatic adult with carcinoma
Death other cause

Model appendix

GENERAL MODEL STRUCTURE

MISCAN-Colon is a stochastic microsimulation model for colorectal cancer (CRC) programmed in Delphi (Borland Software Corporation, Scotts Valley, California, United States). It can be used to explain and predict trends in CRC incidence and mortality and to quantify the effects and costs of primary prevention of CRC, screening for CRC, and surveillance after polypectomy.

The term 'microsimulation' implies that individuals are moved through the model one at a time, rather than as proportions of a cohort. This allows future state transitions to depend on past transitions, giving the model a 'memory'. Furthermore, unlike most traditional Markov models, MISCAN-Colon does not use yearly transition probabilities; instead it generates durations in states, thereby increasing model flexibility and computational performance. The term 'stochastic' implies that the model simulates sequences of events by drawing from distributions of probabilities/ durations, rather than using fixed values. Hence, the results of the model are subject to random variation.

MISCAN-Colon consists of 3 modules: a demography module, natural history module, and screening module.

THE DEMOGRAPHY MODULE

Using birth- and life-tables representative for the population under consideration, MISCAN-Colon draws a date of birth and a date of non-CRC death for each individual simulated. In MISCAN-Colon the maximum age an individual can achieve is exactly 100 years.

THE NATURAL HISTORY MODULE

Transitions

As each simulated person ages, one or more adenomas may develop (Supplemental Figure 1). These adenomas can be either progressive or non-progressive. Both progressive and non-progressive adenomas can grow in size from small (≤5mm), to medium (6-9mm), to large (≥10mm); however, only progressive adenomas can develop into preclinical cancer. A preclinical cancer may progress through stages I to IV; however, during each stage, CRC may be diagnosed because of symptoms. After clinical diagnosis, the survival depends on the stage of the cancer. For individuals with synchronous CRCs

at time of diagnosis, the survival of the most advanced cancer is used. The date of death for individuals with CRC is set to the earliest simulated death (either due to CRC or due to another cause (see: 'The demography module')).

Transition Probabilities and Durations in States

An individual's risk of developing adenomas depends on the individual's age and a personal risk index. As a result of the latter most individuals develop no adenomas, whilst some develop many. We assumed that the distribution of adenomas over the colon and rectum is equal to the distribution of cancers in the Netherlands before the introduction of screening (between 1999 and 2003).[1] The age-specific onset of adenomas and the dispersion of the personal risk index were calibrated to data on the prevalence and multiplicity distribution of adenomas as observed in autopsy studies (Supplemental Figure 2).[2-11] The age-specific probability of adenoma-progressivity and the age- and localization-specific transition probabilities between preclinical cancer stages and between preclinical and clinical cancer stages were simultaneously calibrated to data on the age-, stage-, and localization-specific incidence of CRC in the Netherlands before the introduction of screening (between 1999 and 2003) (Supplemental Figure 3).[1]

The average durations between the preclinical cancer stages were calibrated to the rates of screen-detected and interval cancers observed in randomized controlled trials evaluating screening using guaiac fecal occult blood tests.[12-14] This exercise has been described extensively in a publication by Lansdorp-Vogelaar and colleagues.[15] The average duration from the emergence of an adenoma (state 2) until progression into preclinical cancer (state 7) (i.e. the adenoma dwell-time) was calibrated to the rates of interval cancers (including surveillance detected cancers) observed in a randomized controlled trial evaluating once-only sigmoidoscopy screening (Supplemental Figure 4).[16] We assumed an equal overall dwell-time for adenomas developing into CRC from a medium size (30% of all CRCs) and from a large size (70% of all CRCs). All durations in the adenoma and preclinical cancer phase were drawn from exponential distributions. Durations within the adenoma phase and within the preclinical cancer phase were assumed to be perfectly correlated (i.e. if a small adenoma grows into a medium-sized adenoma rapidly, it will also grow into a large adenoma or develop into CRC rapidly); however, durations in the adenoma phase were assumed to be uncorrelated with durations in the preclinical cancer phase (i.e. a rapidly growing adenoma does not necessarily develop into a rapidly progressing cancer). The proportion of medium sized, non-progressive adenomas growing large and the average duration in duration in the medium size, non-progressive adenoma state (state 5) were calibrated to size-specific adenoma detection rates observed in a Dutch randomized controlled trial on

colonoscopy screening (73% small adenomas, 15% medium sized adenomas, 12% large adenomas).[17]

The Screening Module

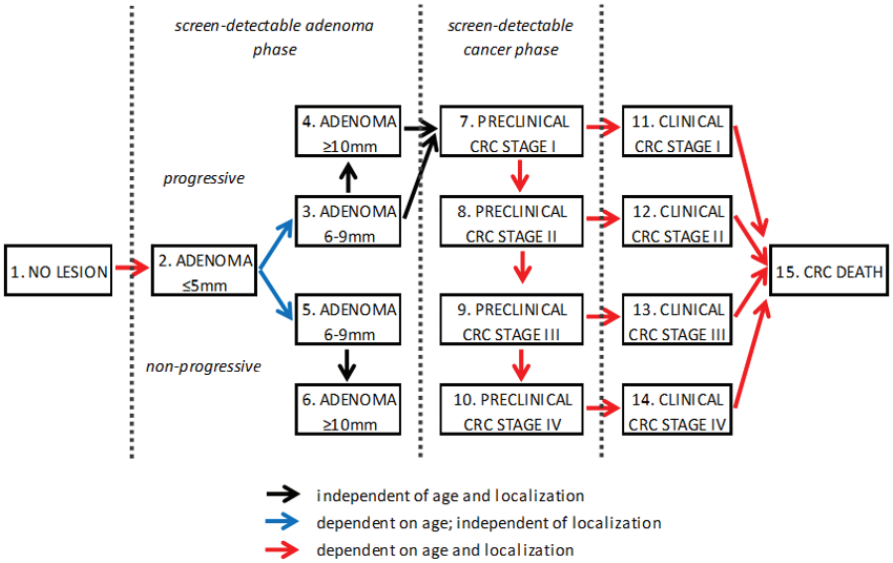
Screening will alter some of the simulated life histories: Some cancers will be prevented by the detection and removal of adenomas; other cancers will be detected in an earlier stage with a more favorable survival. As the stage-specific survival of screen-detected CRC as observed in randomized controlled trials on guaiac fecal occult blood testing was substantially more favorable than that of clinically detected CRC, even after correcting for lead-time bias, we assigned those screen-detected cancers that would have been clinically detected in the same stage the survival corresponding to a one stage less progressive cancer. Hence, a cancer screen-detected in stage II, that would also have been clinically diagnosed in stage II, is assigned the survival of a clinically diagnosed stage I cancer. The only exceptions were screen-detected stage IV cancers. These cancers were always assigned the survival of a clinically diagnosed stage IV cancer.

Besides modeling positive health effects of screening, we also model colonoscopy-related complications and over-diagnosis and over-treatment of CRC (i.e. the detection and treatment of cancers that would not have been diagnosed without screening).

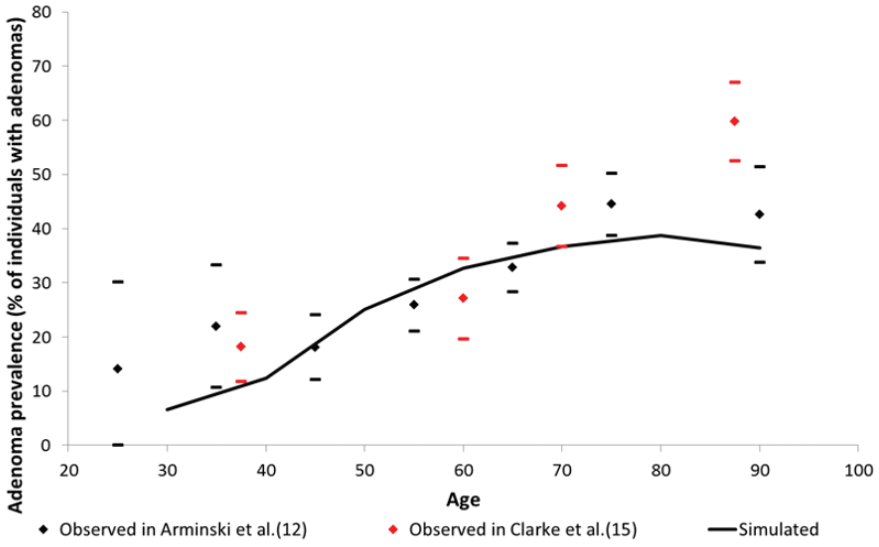
Integrating Modules

The demography module generates a date of birth and a date of non-CRC death for each individual simulated, creating a life-history without adenomas or CRC. In Patient A in Supplemental Figure 5, the natural history module generates an adenoma. This adenoma progresses into preclinical cancer, which is diagnosed because of symptoms in stage II and results in CRC death before non-CRC death would have occurred. In the screening module a screening examination is simulated, indicated by the blue arrow. During this examination the adenoma is detected, and as a result both CRC and CRC death are prevented. Hence, in Patient A, screening prolongs life by the amount indicated by the green arrow. Patient B also develops an adenoma, and although this adenoma does progress into preclinical cancer, Patient B would never have been diagnosed with CRC in a scenario without screening (see *life history 2*). However, during the screening examination simulated in the screening module, again indicated by the blue arrow, CRC is screen-detected in stage I. Hence, in this patient screening results in over-diagnosis of CRC: It detects a cancer that would never have been diagnosed in a scenario without screening. Hence, screening does not prolong life, but it does result in additional LYs with CRC care (over-treatment) as indicated by the red arrow.

Supplemental Figure 1. An Overview of the Natural History Module of MISCAN-Colon.

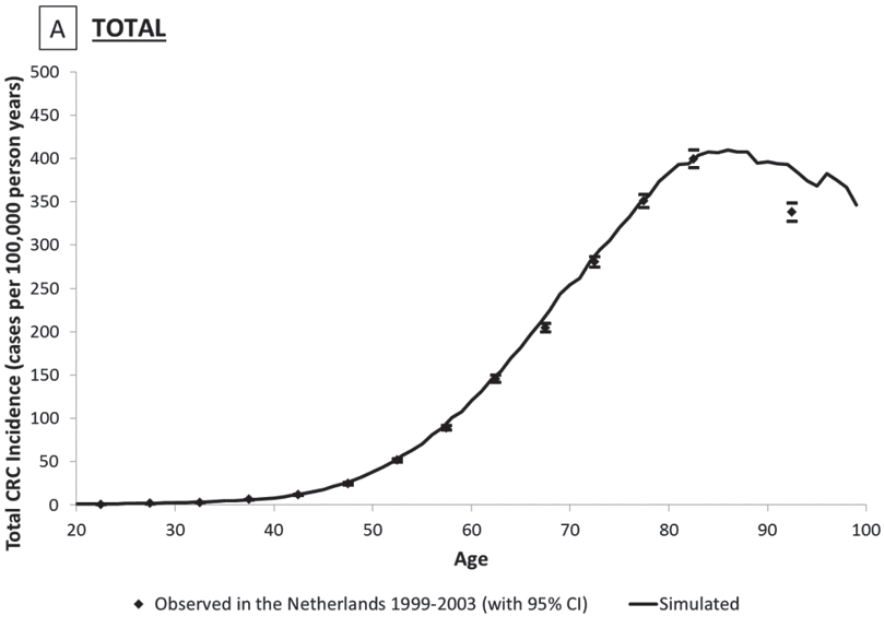


Supplemental Figure 2. Adenomas Prevalence Simulated by MISCAN-Colon Versus Observed in Selected Autopsy Studies and corrected for differences in CRC incidence with the Netherlands.*

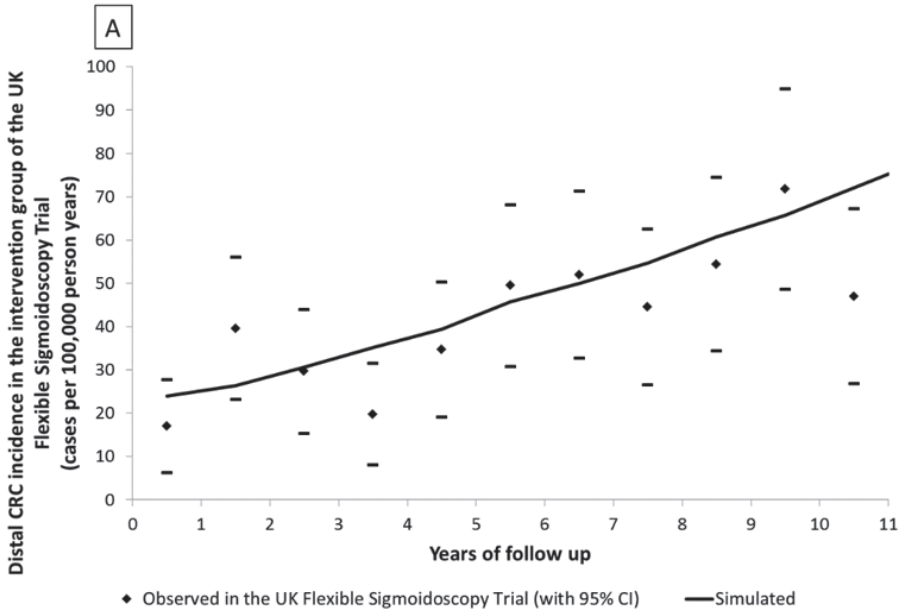


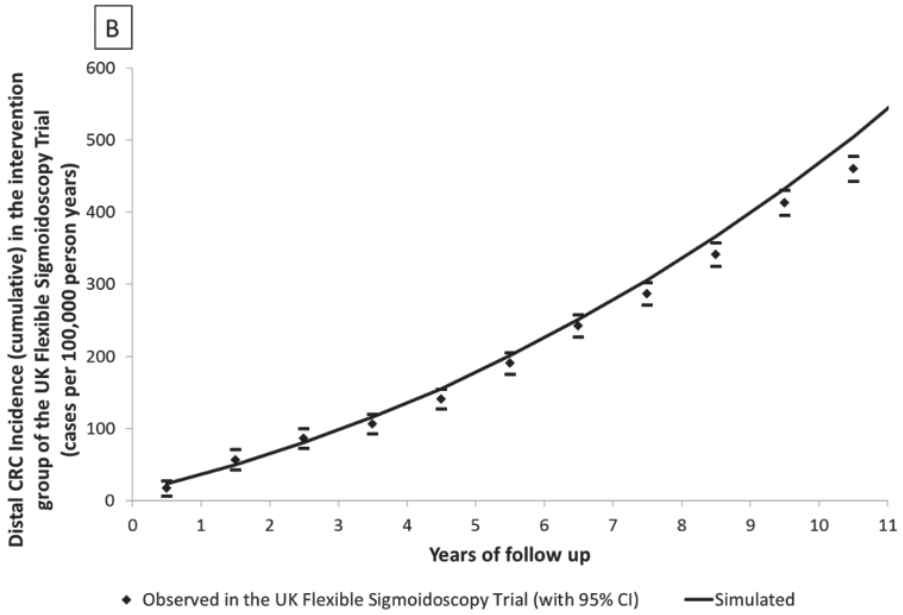
*Observed results are only shown for the two largest studies on which the model has been calibrated. MISCAN-Colon has additionally been calibrated to 8 other autopsy studies.

Supplemental Figure 3. CRC Incidence Observed Before the Introduction of Screening Versus Simulated by MISCAN-Colon; cases per 100,000 person years)

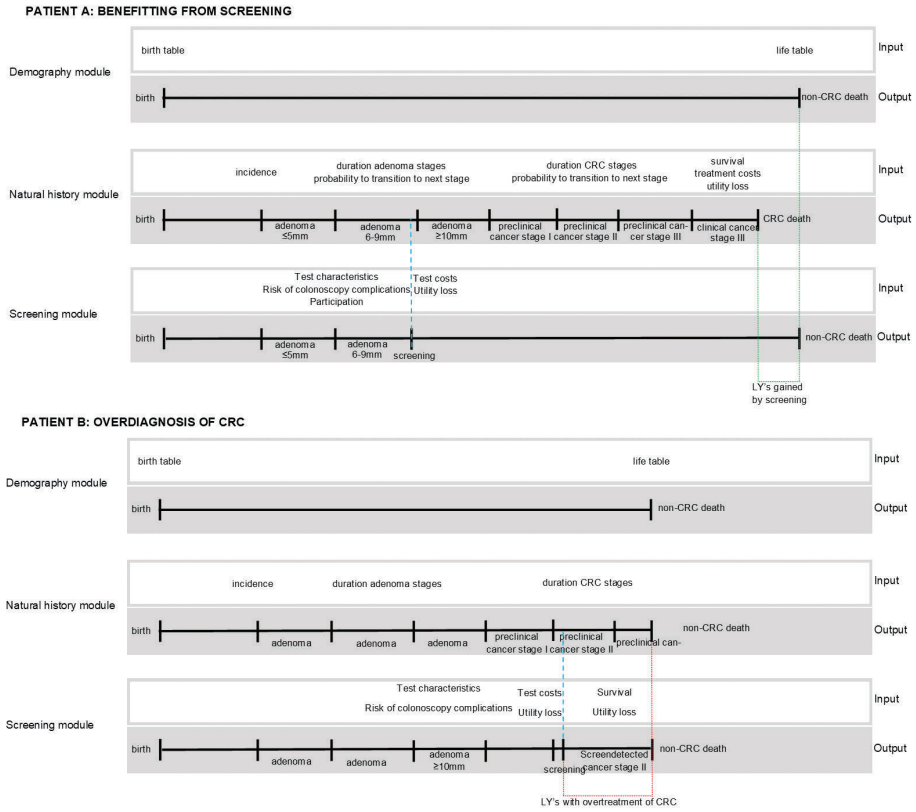


Supplemental Figure 4. Distal CRC Incidence Observed in the Intervention Group of the UK Flexible Sigmoidoscopy Trial Versus Simulated by MISCAN-Colon (per year of follow-up (A), cumulative (B); cases per 100,000 person years).



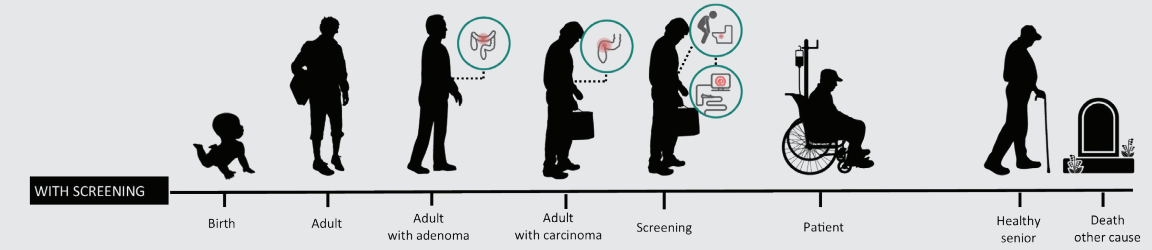


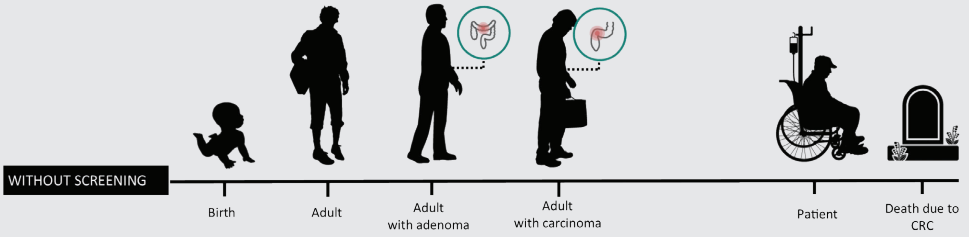
Supplemental Figure 5. Integrating Modules: Two example Patients.



REFERENCES

1. Comprehensive Cancer Centre, vol. 2012: Dutch cancer registry.
2. Arminski TC, McLean DW. Incidence and Distribution of Adenomatous Polyps of the Colon and Rectum Based on 1,000 Autopsy Examinations. *Dis Colon Rectum* 1964;7: 249-61.
3. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. *Cancer* 1988;61: 1472-6.
4. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. *Ann Surg* 1963;157: 223-6.
5. Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, Jensen OM, Koskela E, MacLennan R, Simpson JG, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. *Int J Cancer* 1985;36: 179-86.
6. Jass JR, Young PJ, Robinson EM. Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. *Gut* 1992;33: 1508-14.
7. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. *Scand J Gastroenterol* 1989;24: 799-806.
8. L B. Polyps of the colon and rectum. *Dis Colon Rectum* 1961;4: 249-61.
9. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. *Cancer* 1979;43: 1847-57.
10. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. *Cancer* 1982;49: 819-25.
11. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut* 1982;23: 835-42.
12. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. *Cancer* 1986;58: 397-403.
13. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996;348: 1467-71.
14. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst* 1999;91: 434-7.
15. Lansdorp-Vogelaar I, van Ballegooijen M, Boer R, Zauber A, Habbema JD. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. *Cancer* 2009;115: 2410-9.
16. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, Parkin DM, Wardle J, Duffy SW, Cuzick J, Investigators UKFST. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010;375: 1624-33.
17. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, van de Vijver MJ, Biermann K, Thomeer M, van Leerdam ME, Fockens P, Stoker J, 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: 55-64.


WITH SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Screening
Patient
Healthy senior
Death other cause


WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Patient
Death due to CRC

Summary

Colorectal cancer (CRC) is the second-most common cause of cancer mortality in the western world. The incidence in the Netherlands has been steadily rising from 7,100 cases in 1990 to 13,028 in 2013. Colorectal cancer is believed to develop from a precursor lesion, the adenoma, and more recently an alternative pathway to CRC is described, called the sessile serrated polyp pathway. About 30% of adults between the age of 50 and 75 years old has adenomas in their colorectum. Only a small percentage of these adenomas will eventually develop into a CRC.

With screening, if these adenomas are removed, colorectal cancer can be prevented. If an adenoma has developed into CRC, it mostly does not give symptoms right away and is called preclinical CRC. With screening, CRC can be detected in this preclinical stage. Because treatment will then take place at an earlier stage, there is a higher probability of complete resection, a lower probability of side effects and of lymphatic or distant metastasis. Thus, prevention and early detection of CRC can improve the survival and limit the need for harmful treatments. Indeed, several randomized controlled trials have demonstrated a CRC mortality reduction ranging from 15%-33%.

Colorectal cancer screening is therefore widely adopted across the world. However, the screening programs differ in the way they are organized, the choice of screening test and the age range and interval of screening. The different screening tests already used for CRC screening are the guaiac faecal occult blood test (gFOBT), faecal immunochemical test (FIT), colonoscopy, sigmoidoscopy and computed tomography colonography (CTC).

In the Netherlands, pilot-studies were performed to investigate the acceptance and performance of CRC screening with different screening tests (gFOBT, sigmoidoscopy, FIT, CTC and colonoscopy). FIT was observed to have the highest participation rate. Based on these findings, a population-based CRC screening programme using biennial FIT was introduced in the Netherlands in 2014 with a gradual roll-out period of five years.

In the Dutch screening programme, men and women are screened the same, while men have a higher incidence of CRC than women. In 2013, 7335 men and 5693 women were diagnosed with CRC. The cumulative lifetime incidence of CRC for men is 7.5% versus 6.2% for women. Therefore, in **Chapter 2**, data from one of the Dutch pilot studies were used to show the performance of FIT in men and women. It showed that men and women do not have significantly different positive predictive values for most cut-off levels, resulting in a similar harm-to-benefit ratio. It also showed that men do, however, have higher positivity rates than women, reflected by both higher detection

rates and a higher false positive rate. In **Chapter 3**, the positivity and detection rates of Chapter 2 were used in the MISCAN-Colon model to estimate sex specific FIT sensitivity and specificity. We showed that sensitivity of FIT is lower in women than in men and FIT screening therefore yields less benefit compared to no screening. However, despite the difference in FIT performance and incidence, the incremental costs and benefits of more intensive screening compared to less intensive screening is similar for both sexes. Moreover, the majority of the optimal FIT screening strategies are identical for men and women. Consequently, the QALYs gained for uniform screening and screening stratified by sex never differ more than 7% and are equal for willingness-to-pay thresholds of more than €1300. In other words, there is little to no benefit from FIT screening stratified by sex compared to uniform screening, and our findings support the current policy of uniform FIT screening.

Published mortality rates of colorectal cancer screening have been based on gFOBT screening with suboptimal participation. Long-term effectiveness of FIT screening is expected to be higher because of higher yield and participation and has been estimated using modeling. These models assumed that FIT results are independent of each other over screening rounds. However, if some adenomas do not bleed over several years, they will cause systematic false-negative FIT results. Observed data from two FIT screening rounds of the CORERO study enabled us to further evaluate this issue. In **Chapter 4**, we compared observed second round adenoma detection rates with rates simulated by the MISCAN-Colon model and estimated that over 70% of non-advanced adenomas and 26% of advanced adenomas have to be systematically missed to explain the observed rates. Subsequently, we estimated that incorporating this systematic FIT failure into the model leads to a lower estimate of effectiveness of a FIT screening program. This will impair the effectiveness of a FIT screening program with approximately 6%, thus the effect is limited.

Even though FIT was estimated to result in the highest participation rate in the Dutch pilot studies, the detection rate per participant in a single screening round was significantly higher with CTC and colonoscopy. The long-term cost-effectiveness of both these tests is therefore interesting to estimate. Earlier cost-effectiveness analyses comparing colonoscopy with CT-colonography (CTC) screening showed that the results were highly sensitive to the assumptions regarding attendance and also to test costs. At that time comparative data for these assumptions were lacking. In **Chapter 5**, we used data from the COCOS-trial comparing CTC and colonoscopy in a dedicated screening setting to perform a more representative cost-effectiveness analysis. We used the microsimulation model MISCAN-Colon to compare the Quality Adjusted Life years (QALYs) gained and the costs of several screening strategies with CTC and colonoscopy. Even though colonoscopy was the preferred test if participation would be equal, with

the participation as observed in COCOS, colonoscopy was only more cost-effective in the screening strategies with one or two lifetime screens, whereas CTC was more cost-effective in strategies with more lifetime screens. CTC was the preferred test for willingness-to-pay-thresholds of €3,200 per QALY gained and higher, which is lower than the Dutch willingness-to-pay threshold of €20,000. Therefore, CTC screening is more cost-effective than colonoscopy screening. However, the implementation of CTC screening requires prior satisfactory resolvement of how to deal with extracolonic findings.

The incidence of CRC does not differ for different levels of socioeconomic status (SES) in the Netherlands. The mortality due to CRC, however, does differ: individuals with a lower SES (more deprivation, lower income, lower level of education) have more chance to ever die from CRC. In addition, participation in other screening programs around the world is most often lower in more deprived areas. Therefore, the implementation of population screening could influence these health inequalities. In **Chapter 6**, we investigated SES differences in participation and diagnostic yield of the Dutch FIT screening programme in 2014 and 2015. 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, detection rate and yield across SES quintiles. 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). 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%, 95%CI 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%). Because of a higher detection in more deprived participants, screening has the potential to reduce health inequalities in CRC mortality. However, this is currently offset by the lower participation in this group.

Individuals with adenomas are at increased risk of developing metachronous adenomas and CRC, even after the adenomas have been completely removed. Therefore, surveillance is recommended for individuals that have had a polypectomy. Prior research shows that surveillance is currently often not used efficiently. In the Netherlands, the compliance to the guideline was also reported to be low. This low compliance to the Dutch guideline for colonoscopy surveillance after polypectomy led to the release of a new guideline in 2013. This new guideline was risk-stratified at a more detailed level than the previous one to achieve more efficient use of colonoscopy resources. In **Chapter 7**, we assessed the feasibility of the risk-stratified guideline by evaluating the correct interpretation of and adherence to this guideline. In an online survey consisting of 15 example cases of patients with polyps, gastroenterologist gave

a correct recommendation in a median of 10 out of 15 cases. The percentage of correct recommendations per case ranged from 14% to 95% (median case 76%), which is higher than in a previously published survey on the adherence to the older guideline. If their recommended intervals deviated from the new guideline, respondents were asked to indicate their motives for doing so. Deviations were mainly due to misinterpretation of the guideline with respect to serrated polyps (48%) or misreading of the questions (30%). Chapter 7 shows that detailed (more complex) risk stratification for designation of a surveillance interval is feasible. Compliance could potentially be further improved by clarifying the correct way to handle specific aspects of the guideline.

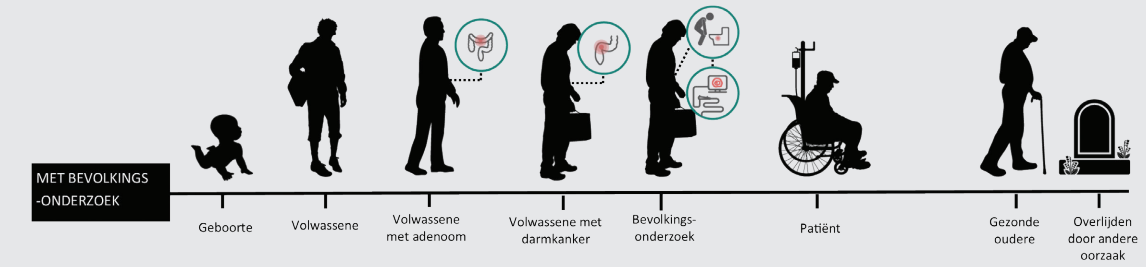
CONCLUSIONS

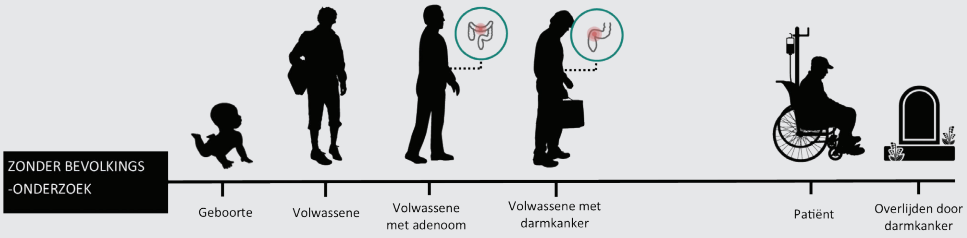
- Men have higher positivity rates than women, reflected by both higher detection rates and a higher false positive rate. (Chapter 2)
- Despite sex differences in FIT performance and CRC risk, CRC screening stratified by sex does not improve cost-effectiveness. (Chapter 3)
- A comparison between observed and modeled second round detection rates shows that a proportion of adenomas is systematically missed by repeated FIT screening. This impairs the efficacy of FIT screening slightly. (Chapter 4)
- Based on randomized controlled screening trial data, showing higher participation rates for CTC, CTC is more cost-effective than colonoscopy screening for colorectal cancer. (Chapter 5)
- 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 program, this is currently offset by the lower participation in this group. (Chapter 6)
- The adherence to the updated risk-stratified colonoscopy surveillance guideline seems reasonable, and is higher than the adherence to the previous simple guideline. A detailed (more complex) risk stratification for designation of a surveillance interval is therefore feasible. (Chapter 7)

RECOMMENDATIONS

- In case of sufficient colonoscopy capacity, FIT screening should not be stratified by sex. (Chapter 2 and 3)
- The program performance of FIT screening over multiple screening rounds should be estimated in future studies to validate the dependency of FIT test results. (Chapter 4)
- CTC should be included in future studies that evaluate alternative screening tests for the FIT. (Chapter 5)

- Further research is needed to clarify on how to target socioeconomic groups that are less compliant and/or more at risk for CRC and ensure well-informed decision-making. (Chapter 6)
- To improve adherence to the guideline for surveillance after polypectomy, clarification on how to score serrated adenomas and high-grade dysplasia should be given, and the use of a pocket card and/or app should be further encouraged. (Chapter 7)


MET BEVOLKINGS
-ONDERZOEK
Geboorte
Volwassene
Volwassene
met adenoom
Volwassene met
darmkanker
Bevolkings-
onderzoek
Patiënt
Gezonde
oudere
Overlijden
door andere
oorzaak


ZONDER BEVOLKINGS
-ONDERZOEK
Geboorte
Volwassene
Volwassene
met adenoom
Volwassene met
darmkanker
Patiënt
Overlijden door
darmkanker

Samenvatting

Dikke darmkanker, of kortweg darmkanker, is de op één na dodelijkste vorm van kanker in het Westen. Het aantal darmkanker gevallen toont een stijgende lijn, van 7.100 gevallen in 1990 tot 13.028 gevallen in 2013. In eerste instantie dacht men dat darmkanker zich voornamelijk ontwikkelt vanuit een voorstadium, het adenoom. Naar schatting heeft 30% van de volwassenen tussen de 50 en 75 jaar adenomen in hun dikke darm. Slechts een klein percentage van deze adenomen zal uiteindelijk uitgroeien tot darmkanker. Sinds kort is bekend dat darmkanker zich ook kan ontwikkelen vanuit een sessiel serrated poliep.

Als adenomen door screening verwijderd worden, kan darmkanker voorkomen worden. Indien een adenoom zich ontwikkelt tot darmkanker, leidt dit meestal niet direct tot symptomen, dit wordt pre-klinische darmkanker genoemd. Met screening kan darmkanker gedetecteerd worden in dit pre-klinische stadium. Omdat in dat geval de behandeling in een eerder stadium gestart kan worden, is er een hogere kans op volledig herstel, een lagere kans op bijwerkingen en ook een lagere kans op uitzaaiingen naar de lymfeklieren of andere organen. Kortom, vroege detectie van darmkanker kan de overlevingskans verbeteren, de noodzaak aan schadelijke behandelingen beperken, en zelfs darmkanker voorkomen. Verschillende gerandomiseerde studies hebben uitgewezen dat screening het aantal doden aan darmkanker kan terugdringen met 15-33%.

Vanwege bovengenoemde voordelen wordt darmkanker screening wereldwijd toegepast; de verschillende screening programma's varieren echter sterk in de manier van organisatie, het type test, start- en stopleeftijd en screeningsinterval. De verschillende testen die gebruikt worden voor darmkanker screening zijn de guiac fecale occult bloedtest (gFOBT), de fecale immunochemische test (FIT), coloscopie, sigmoïdoscopie en computertomografie colonografie (CTC).

In Nederland zijn proefbevolkingsonderzoeken uitgevoerd om de acceptatie en het presteren van de verschillende screeningtesten te onderzoeken (gFOBT, sigmoidoscopy, FIT, CTC en colosopie). Deze proefbevolkingsonderzoeken lieten zien dat de meeste mensen meededen aan screening met FIT. Gebaseerd op deze bevindingen, is in 2014 een georganiseerd landelijk bevolkingsonderzoek naar darmkanker geïntroduceerd met een tweejaarlijkse FIT en een geleidelijke uitrol periode van vijf jaar.

In het Nederlandse bevolkingsonderzoek worden mannen en vrouwen op gelijke wijze gescreend, terwijl mannen een hoger risico op darmkanker hebben dan vrouwen. Zo is in 2013 bij 7335 mannen en 5693 vrouwen de diagnose darmkanker vastgesteld en bedraagt het cumulatief risico op darmkanker tijdens het gehele leven 7,5% voor

mannen, tegenover 6,2% voor vrouwen. In **hoofdstuk 2** van dit proefschrift zijn met behulp van data van één van de Nederlandse proefbevolkingsonderzoeken de prestatie van FIT in mannen en vrouwen onderzocht. Dit hoofdstuk laat zien dat FIT geen significant verschil heeft voor mannen en vrouwen op positieve voorspellende waarde op de meeste afkapwaarden. Daarnaast blijkt dat mannen vaker een positieve test hebben (verwijspercentage) dan vrouwen, wat wordt veroorzaakt doordat er meer adenomen en darmkanker worden gevonden (detectiecijfer) en omdat ze vaker een onterecht positieve test hebben.

In **hoofdstuk 3** zijn de positiviteits- en detectiecijfers van hoofdstuk 2 gebruikt in het MISCAN-Colon model om geslachtsspecifieke sensitiviteit en specificiteit van FIT te berekenen. Dit hoofdstuk laat zien dat de sensitiviteit voor FIT lager is voor vrouwen dan voor mannen en dat FIT screening daarom minder voordeel heeft bij hen ten opzichte van geen screening. Desondanks waren de incrementele kosten en effecten van meer intensieve screening vergeleken met minder intensieve screening gelijk voor beide geslachten. Ook is de meerderheid van de optimale FIT screeningsstrategieën gelijk voor mannen en vrouwen. Vanaf een betalingsbereidheid van meer dan €1300 per gewonnen QALY is er geen verschil tussen uniforme screening en screening gestratificeerd naar geslacht. Met andere woorden, er is weinig tot geen voordeel te behalen met het stratificeren van FIT screening naar geslacht vergeleken met uniforme screening, en dus ondersteunen deze bevindingen het huidige beleid van uniforme FIT screening.

Voor gFOBT en sigmoïdoscopie zijn gerandomiseerde studies uitgevoerd met een geobserveerde afname in sterfte aan darmkanker door screening. Er zijn geen gerandomiseerde studies met FIT uitgevoerd en daarom wordt de afname in sterfte aan darmkanker door FIT screening geschat met behulp van modellen. Deze modellen nemen aan dat FIT-resultaten van de verschillende ronden onderling onafhankelijk zijn. Echter, gebleken is dat, als adenomen gedurende meerdere jaren niet bloeden, dit leidt tot een systematische fout negatieve FIT-resultaten. Geobserveerde data van twee FIT-screeningrondes van de CORERO-studie stelden ons in de gelegenheid om dit probleem verder te evalueren.

In **hoofdstuk 4** is het geobserveerde detectiecijfer van adenomen in de tweede ronde van deze studie vergeleken met het detectiecijfer van adenomen zoals gesimuleerd door het MISCAN-Colon model. Op basis hiervan schatten we dat 70% van de non-advanced adenomen en 26% van de advanced adenomen systematisch gemist zou moeten worden om de geobserveerde detectiecijfers te verklaren. Vervolgens hebben we deze systematische testresultaten in het model gebruikt en berekend dat het systematisch missen van adenomen leidt tot een beperkte afname in de effectiviteit van het FIT-screeningprogramma van ongeveer 6%.

Hoewel de deelnamegraad in de proefbevolkingsonderzoeken het hoogste was voor de FIT, met CTC en coloscopie werden sgifnicant meer relevante afwijkingen gevonden in één screeningsronde dan met FIT. Daarom hebben we de lange termijn kosteneffectiviteit van CTC en coloscopie onderzocht. Eerdere kosteneffectiviteitsanalyses die beide tests vergeleken toonden aan dat de resultaten erg gevoelig zijn voor aannames over deelname en kosten van de test. Destijds ontbraken data voor deze aannames in een vergelijkbare setting.

In **hoofdstuk 5** zijn data uit de COCOS trial gebruikt, waar CTC werd vergeleken met coloscopie in een toegewijde screening setting, om een meer representatieve kosteneffectiviteitsanalyse uit te voeren. Met behulp van het microsimulatie model MISCAN-Colon zijn de gewonnen QALY's en de kosten van verschillende screeningstrategieën vergeleken voor CTC en coloscopie.

Ook al was coloscopie de voorkeurstest bij een gelijke deelnamegraad, met de deelnamegraad zoals gemeten in de COCOS-trial was coloscopie alleen kosteneffectiever in strategieën met één of twee screeningsrondes. Vanaf een betalingsbereidheid van €3,200 per gewonnen QALY was CTC de voorkeursscreeningstest. Omdat de Nederlandse kostenbereidheid €20,000 per gewonnen QALY is, beschouwen we CTC-screening als kosteneffectiever dan screening met coloscopie. Echter, voordat CTC screening geïmplementeerd kan worden, zal eerst een bevredigende oplossing gevonden moeten worden voor de extracolonische bevindingen.

Het risico op darmkanker is niet afhankelijk van sociaal economische status (SES) in Nederland, maar de kans om aan darmkanker te overlijden wel. Individuen met een lagere SES (meer deprivatie, lager inkomen, lager opleidingsniveau) hebben een hogere kans op overlijden, en uit voorgaand onderzoek weten we dat de deelname aan bevolkingsonderzoeken lager is voor individuen met een lage SES. Het bevolkingsonderzoek naar darmkanker zou de gezondheidsverschillen tussen verschillende SES daarom verder kunnen vergroten.

In **hoofdstuk 6** hebben we dit verder onderzocht, door de SES-verschillen in deelnamegraad en diagnostische opbrengst in het Nederlandse bevolkingsonderzoek met FIT in 2014 en 2015 te vergelijken. De SES gebaseerd op postcode werd gebruikt als maat voor SES en alle postcodes werden verdeeld in vijf groepen (kwintielen), met in kwintiel 1 de postcodes met de hoogste SES en in kwintiel 5 die met de laagste SES. We hebben vervolgens met behulp van logistische regressie de deelnamegraad, het detectiecijfer en de opbrengst van het bevolkingsonderzoek van de SES kwintielen vergeleken. De deelname aan FIT-screening was significant lager voor kwintiel 5 (67.0%) vergeleken met de andere kwintielen (73.0% to 75.1%). Het detectiecijfer per deelnemer voor advanced adenomen en darmkanker samen steeg geleidelijk van 3,3% in kwintiel 1 tot 4,0% in kwintiel 5 (adjusted OR 1.20%, 95%CI 1.16-1.24). Door de lagere

deelnamegraad was de opbrengst per genodigde vergelijkbaar voor kwintiel 5 (2,04%) en kwintiel 1 (2.00%), welke beiden lager waren dan kwintiel 2 tot 4 (2.20%-2.28%). Door het hogere detectiecijfer in deelnemers met lage SES heeft darmkankerscreening dus de potentie om gezondheidsverschillen in darmkanker mortaliteit te verkleinen tussen SES groepen. Deze potentie wordt momenteel echter teniet gedaan door de lagere deelname in deze groep.

Individuen met adenomen hebben een verhoogd risico op het ontwikkelen van metachrone adenomen en darmkanker, zelfs nadat de adenomen volledig verwijderd zijn. Daarom wordt surveillance aangeraden voor individuen die een poliepectomie hebben ondergaan. Eerder onderzoek toonde aan dat surveillance op dit moment niet efficiënt wordt ingezet. In Nederlands was de naleving van de richtlijn ook laag en daarom is er een nieuwe richtlijn voor surveillance na poliepectomie uitgebracht in 2013. In overeenstemming met de meest recente inzichten wat betreft risicofactoren voor metachrone adenomen en kanker heeft deze nieuwe richtlijn een gedetailleerdere risicostratificatie dan de vorige richtlijn.

In **hoofdstuk 7** hebben we de uitvoerbaarheid van deze richtlijn met risicostratificatie getoetst door de correcte interpretatie van en de naleving van deze richtlijn te evalueren. In een online vragenlijst, die bestond uit 15 voorbeeldcasussen van patiënten met poliepen (adenomen of sessiel serrated poliepen), gaven MDL-artsen een aanbeveling die in overeenstemming was met de richtlijn in een mediaan van 10 van de 15 casussen. De naleving per casus varieerde van 14% tot 95% (mediane casus 76%); dit is hoger dan een soortgelijk uitgevoerde studie naar de naleving van de eerdere richtlijn. Als de MDL-artsen een aanbeveling gaven die niet in overeenstemming was met de richtlijn, hebben we motieven hiervoor gevraagd. Afwijken van de richtlijn werd vooral veroorzaakt doordat de respondenten de richtlijn verkeerd hadden geïnterpreteerd voor serrated poliepen (48%) en/of omdat ze de vraag verkeerd hadden gelezen (30%). Hoofdstuk 7 laat zien dat een gedetailleerde (complexe) risicostratificatie voor het toekennen van een surveillance interval uitvoerbaar is. De naleving kan mogelijk nog verder verbeterd worden door het scoren van specifieke aspecten in de richtlijn te verduidelijken.

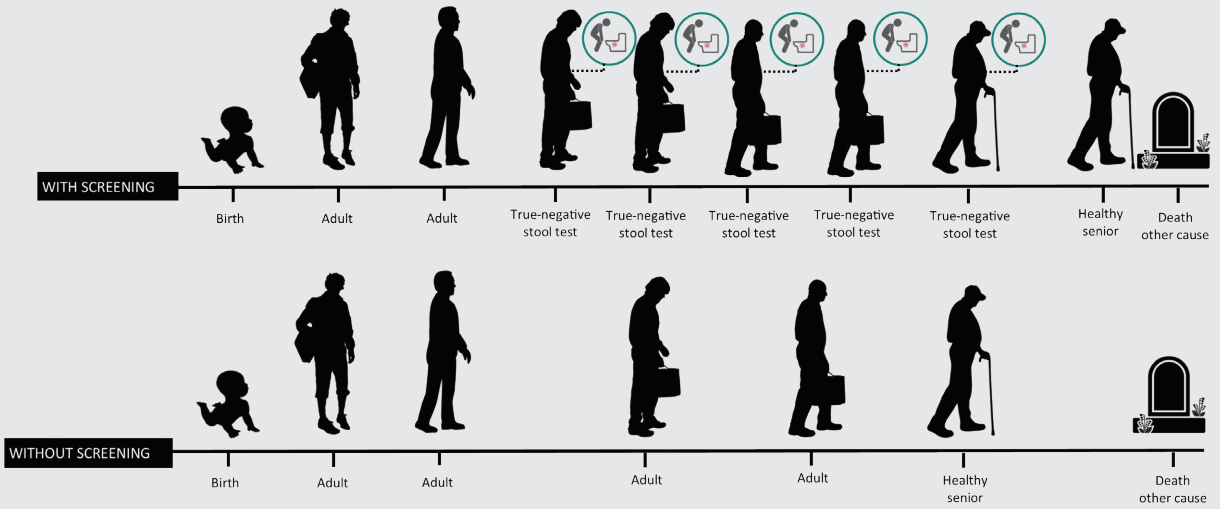
CONCLUSIES

- De FIT heeft bij mannen een hoger verwijscijfer dan vrouwen, veroorzaakt door hogere detectiecijfers en door een hoger percentage fout-positieve testen. (hoofdstuk 2)
- Ondanks verschillen in risico en de prestaties van FIT tussen mannen en vrouwen, is de kosteneffectiviteit van screening op basis van geslacht niet beter dan die van uniforme screening voor iedereen. (hoofdstuk 3)

- Een vergelijking tussen het geobserveerde en gesimuleerde detectiecijfer in tweede ronde FIT-screening, laat zien dat een deel van de adenomen systematisch gemist wordt bij herhaalde FIT-screening. Het systematisch missen van adenomen zorgt voor een lichte daling van de gemodeleerde effectiviteit van FIT-screening ten opzichte van de oorspronkelijk berekeningen. (hoofdstuk 4)
- Gebaseerd op data van een gerandomiseerde studie die een hogere opkomst liet zien voor screening met CTC, is CTC kosteneffectiever dan coloscopie voor darmkanker screening. (hoofdstuk 5)
- Screening heeft de potentie om sociaal-economische gezondheidverschillen in darmkankersterfte te verminderen vanwege een hoger detectiecijfer in deelnemers met een lagere SES. Dit wordt op dit moment echter teniet gedaan door de lagere deelnamegraad in deze groep. (hoofdstuk 6)
- De naleving van de nieuwe coloscopie surveillance richtlijn met risicostratificatie lijkt redelijk en is hoger dan de naleving van de vorige, simpelere richtlijn. Een gedetaillerde, complexere risicostratificatie voor het toekennen van een surveillance interval is dus uitvoerbaar. (hoofdstuk 7)

AANBEVELINGEN

- Als er voldoende coloscopiecapaciteit is, dan moet FIT-screening niet gestratificeerd naar geslacht. (hoofdstuk 2 en 3)
- De prestatie van FIT over meerdere rondes zal bepaald moeten worden in toekomstig onderzoek om de afhankelijkheid van FIT-testresultaten te valideren. (hoofdstuk 4)
- In eventueel toekomstig onderzoek naar een alternatief voor FIT moet ook CTC als alternatief worden meegenomen. (hoofdstuk 5)
- Meer onderzoek is nodig om te bepalen hoe bepaalde sociaal economische groepen met een lagere deelnamegraad en/of een hoger risico op darmkanker benaderd kunnen worden, waarbij de geïnformeerde keuze gewaarborgd wordt. (hoofdstuk 6)
- Om de naleving van de richtlijn voor surveillance na poliepectomie te verbeteren, moet de manier van scoren van serrated poliepen en hooggradige dysplasie worden verduidelijkt. Daarnaast moet het gebruik van een zakkaart en/of app verder worden worden gestimuleerd. (hoofdstuk 7)


WITH SCREENING
Birth
Adult
Adult
True-negative stool test
True-negative stool test
True-negative stool test
True-negative stool test
True-negative stool test
Healthy senior
Death other cause
WITHOUT SCREENING
Birth
Adult
Adult
Adult
Adult
Healthy senior
Death other cause

Dankwoord

Het is toch nog gebeurd, mijn proefschrift is af! Dat het zo lang duurde lag met name aan mijzelf en niet aan de ondersteuning die ik heb gekregen en hoe leuk het werk was. Terwijl ik de laatste jaren niet bepaald enthousiast praatte over de afronding van mijn proefschrift, bleven velen van jullie nog altijd geïnteresseerd naar mijn proefschrift vragen. Voor die ondersteuning en interesse wil ik jullie graag bedanken!

Als eerste wil ik graag mijn co-promotor, prof dr. Harry de Koning en promotor dr. Iris Lansdorp-Vogelaar bedanken. Beste Harry, jij was als promotor alleen via de zijlijn betrokken bij het inhoudelijke gedeelte van mijn proefschrift. Bedankt dat je, toen ik bij je kwam om over de eerste volledige versie van mijn introductie en discussie te praten, meteen de doorzette om naar een datum uit te kijken, dat zetje kon ik goed gebruiken. Beste Iris, ik heb het altijd heel fijn gevonden dat jij mijn directe leidinggevende was, we zaten inhoudelijk altijd snel op één lijn en jij gaf mij de sturing om het project/artikel daarop verder te brengen. Daarnaast heb ik veel gehad aan je constructieve feedback. Bedankt hiervoor en voor je betrokkenheid en lange adem.

Ook wil ik graag de kleine commissie, prof dr Evelien Dekker, prof dr Anton Kunst en dr. Manon Spaander en de leden van de grote commissie bedanken voor de tijd en aandacht die ze aan dit proefschrift hebben besteed.

De eerste jaren van mijn promotietaject was dr. Marjolein van Ballegooijen mijn co-promotor. Beste Marjolein, bedankt voor je begeleiding in die eerste jaren, voor je verrassende inzichten en doordenken tijdens de overleggen, daardoor heb ik geleerd om de onderwerpen van allerlei kanten te bekijken.

Colon groep: Luuk, Frank, Else , Esther, Maaike, Reinier, Sonja. Bedankt voor de prettige samenwerking. Luuk, volgens mij heb ik de eerste jaren behoorlijk vaak bij je aangeklopt, bedankt dat je zo laagdrempelig aanspreekbaar was en voor je begeleiding in het wegwijs worden in MISCAN-Colon. Esther, bedankt dat ik in deze afrondende fase altijd met vragen bij je terecht kon.

Ik heb het geluk gehad om tijdens mijn werk op MGZ altijd op de leukste kamers van de afdeling te hebben gezeten. Ingrid en Daan, bedankt voor alle gesprekken over de belangrijkste roddels over Jelena en sizzurp tussen al het computer- en cijferwerk door. Daardoor heb ik nu alleen wel een nieuwe voorwaarde voor werk: er moet in ieder geval één persoon op de hoogte zijn van de achterklap. In het nieuwe Na gebouw kwamen er weer nieuwe kamergenoten bij. Steffie, Hilde, Timor, Roelie en Tjeerd: bedankt voor alle gezellige lunches, tussendoorpraatjes en uitstapjes. Dat we als kamer een goed team waren werd bevestigd door onze scherpe ontsnappingstijden uit de Escape rooms.

In de tussentijd ben ik op andere leuke plekken gaan werken, en moest in eigen tijd mijn proefschrift afronden. Sake, ik heb daarbij altijd het gevoel gehad dat ik ook ruimte voor mijn promotie mocht nemen, bedankt daarvoor. Dit is ook zo in mijn huidige werk, met name bij THINC., bedankt voor de flexibiliteit en de inspirerende werkomgeving. Of het nu aan de vriendelijke verzoeken tot afronding, een grapje over het lange traject, de ervaringen met kort-cyclisch werken of de positieve sfeer heeft gelegen, het werken bij jullie heeft er aan bijgedragen dat ik over de drempel tot afronding ben heengestapt.

Dan mijn vriendinnen, Diana, Annemieke en alle Moorkopjes: Djoelan, Leonie, Helma, Sara, Henriette, Elsabeth, Kirsten, Titia, Margreet. Bedankt voor het fijne contact dat we nog steeds hebben op onder andere onze plomaweekenden, Sinterklaasvieringen in de lente en op dit moment onze plozo's. Toch verlang ik wel terug naar de tijd dat we zomaar bij elkaar op de thee of borrel konden, dus ook al grappen we er vaak over, ik zou serieus graag later in dat mooie bejaardenoptrekje voor ons allen zitten. Leonie en Helma, jullie leven had in verschillende periodes het meest gemeen met die van mij, met andere redenen, maar beiden bedankt voor het fijne contact in deze periodes. Djoelan, bij jou kan ik altijd aankloppen voor een belletje of een kopje thee, bedankt voor je vrolijke noot. Titia, bedankt voor de input voor mijn Nederlandse samenvatting.

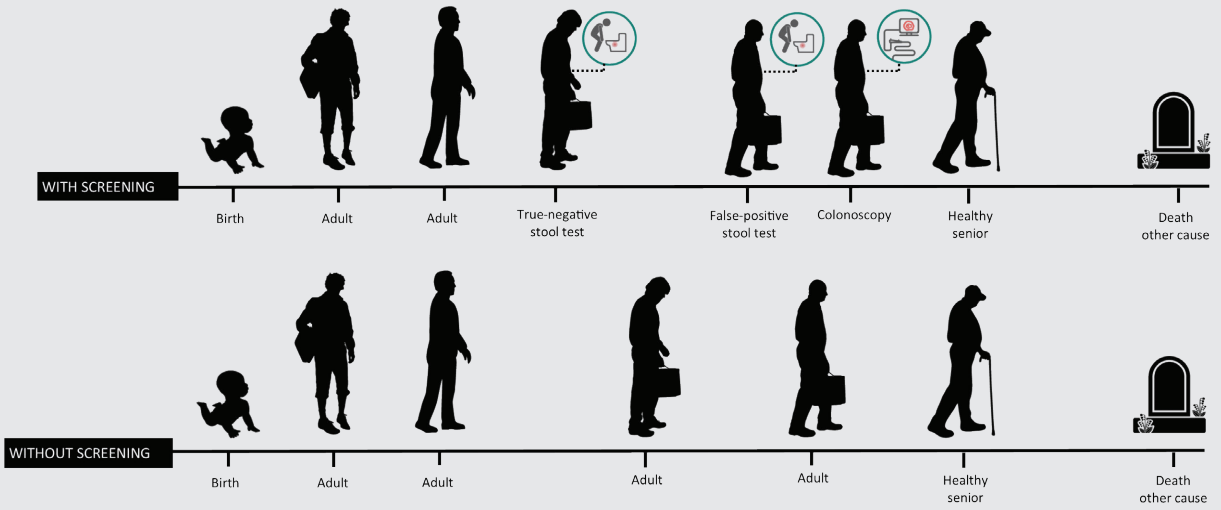
Lieve Rinske en Esther, sada tay! Bedankt dat jullie mijn paranimfen zijn. Al is de invulling wel iets anders dan dit normaal, ik heb nu al veel aan jullie rol gehad, bedankt dat jullie nu mijn vraagbaak zijn voor de promotie en andere kwesties. Maar natuurlijk juist ook bedankt voor de lange vriendschap, ik hoop dat we die voor altijd behouden en als bejaarden nog steeds op No Diggity dansen.

Dan mijn familie: lieve Jan, Ellis en Arjen, bedankt dat jullie altijd voor me klaarstaan, ik hou van jullie! Ik wil ook mijn schoonfamilie bedanken, Dirk en Anjolan, bedankt voor jullie betrokkenheid, Steven, bedankt voor de supergave cover.

Lieve Jarig, op een moment tijdens de co-schappen vroeg je op een aangrijpende manier aan mij of ik later alsjeblieft onderzoek kon doen om de situatie van anderen zoals jou (na een CVA) te verbeteren. Op dat moment wilde ik al liever onderzoek doen in plaats van de praktijk in en voelde ik mij schuldig dat ik niet nuttig werk in de "frontlinie" zou gaan doen. Die vraag van jou is daarom heel belangrijk voor me geweest. Ook al ben je er niet meer, bedankt voor dit inzicht.

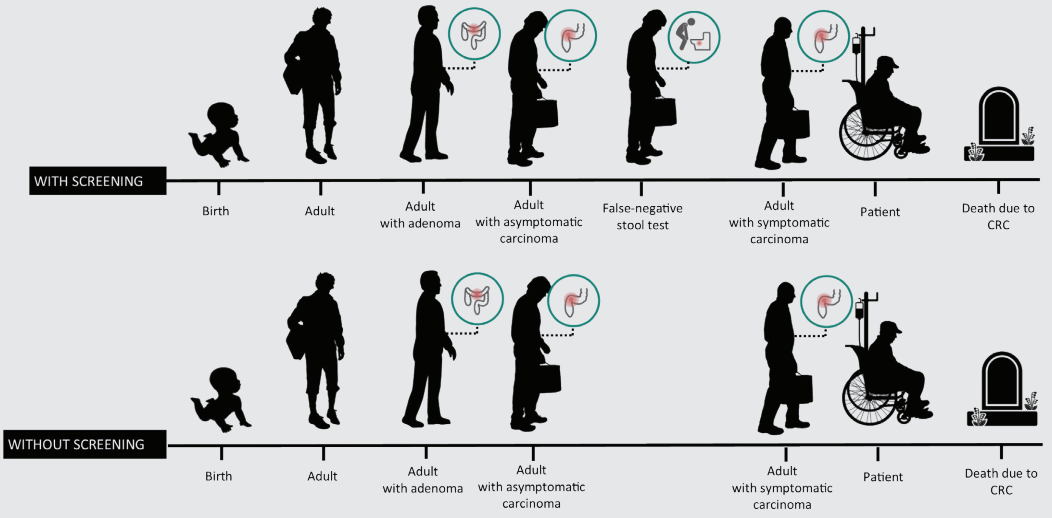
Lieve Pieter, van studerend met mooie lange reizen, gingen we na onze wereldreis beiden het werkende leven in. Dat was wel een overgang, die we interessant hebben weten te houden door ons in avonturen op eigen grond te storten met onze klushuizen. We hebben het geluk gehad dat we Norah en Milou hebben kunnen krijgen, waardoor we het wel nog drukker kregen. Nu mijn promotie is afgerond, de kinderen zelfstandiger worden en het huis steeds idyllischer wordt, komen we in rustiger vaarwater, ik ben benieuwd of we daar van gaan genieten of weer een nieuw avontuur aangaan.

Lieve Norah en Milou, ik ben heel blij met jullie. Ook al hebben jullie niet bijgedragen aan de snelheid waarmee ik dit proefschrift kon afronden, jullie zorgen voor afleiding én relativering: er zijn belangrijkere dingen in het leven dan werk. Ik hou van jullie!


WITH SCREENING
Birth
Adult
Adult
True-negative stool test
False-positive stool test
Colonoscopy
Healthy senior
Death other cause
WITHOUT SCREENING
Birth
Adult
Adult
Adult
Adult
Healthy senior
Death other cause

Curriculum vitae

Miriam Petra van der Meulen was born on February 1st 1984 in Leeuwarden, the Netherlands. She completed her secondary school in 2002 at the Piter Jelles Montessori in Leeuwarden. In the same year, she started studying Nursing at the Noordelijke Hogeschool Leeuwarden (NHL). In 2003, she started studying medicine at the RIjksuniversiteit Groningen (RuG) and in 2011, she received her Master of Science degree. She completed her final internship at the Prince Wellington Ngu Hospital in Vava'u in Tonga and wrote her master thesis on the effect of life events on the morning response of cortisol in adolescents at the Interfacultair Centrum voor Psychiatrische Epidemiologie, Universitair Medisch Centrum Groningen (UMCG). In 2011, she started to work at the department of Public Health at the Erasmus MC on colorectal cancer screening. During her time at the department of Public Health, she obtained her master degree Health sciences with specialization Public Health at the Netherlands Institute of Health Sciences (NIHES). In 2017, she started to work at the same department as a researcher infectious diseases modeling. Currently, Miriam works as an HTA researcher for the Julius Center and The Healthcare Innovation Centre (THINC.) at UMC Utrecht.

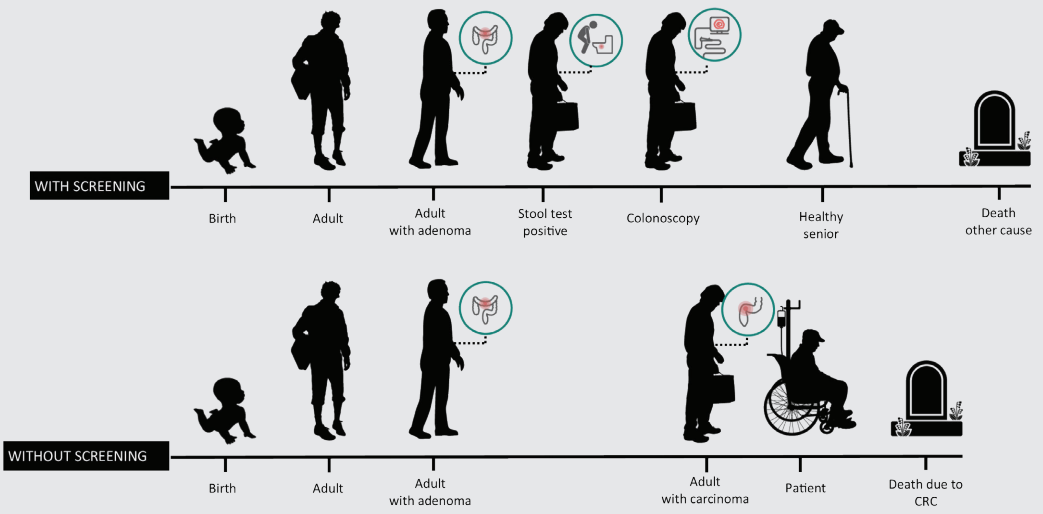

WITH SCREENING
Birth
Adult
Adult with adenoma
Adult with asymptomatic carcinoma
False-negative stool test
Adult with symptomatic carcinoma
Patient
Death due to CRC
WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with asymptomatic carcinoma
Adult with symptomatic carcinoma
Patient
Death due to CRC

PhD portfolio

Name PhD student:	Miriam Petra van der Meulen
Erasmus MC department:	Public Health
Research School:	Netherlands Institute for Health Sciences (NIHES)
PhD period:	2011-2020
Promotor:	Prof. dr. H.J. de Koning
Co-promotor:	Dr. I. Lansdorp-Vogelaar

PhD training	Year	Workload
General academic skills		
Internal writing group department of public health, Erasmus MC	2011	40 hours
Internal writing group department of public health, Erasmus MC	2014	40 hours
Course, planning and Evaluation of screening, NIHES	2011	40 hours
Master of Health sciences, specialisation Public Health, NIHES	2011-2015	70 ECTS*
Presentations		
Cancer Intervention and Surveillance Modeling Network (CISNET) meeting, Bethesda Washington	2011	20 hours
United European Gastroenterology Week (UEGW), Amsterdam, the Netherlands (poster)	2012	20 hours
Digestive Disease Week (DDW), Orlando, Florida (poster)	2013	20 hours
NVGE, Veldhoven, the Netherlands	2013	40 hours
Research seminar at department of public health, Erasmus MC	2014	20 hours
United European Gastroenterology (UEG) Workshop, Brno, Poland	2015	40 hours
International Cancer Screening Network (ICSN), Rotterdam, the Netherlands	2015	20 hours
Attendance to (inter)national conferences		
ESDO GI Cancer Workshop Stockholm	2011	40 hours
Cancer Intervention and Surveillance Modeling Network (CISNET) meeting, Bethesda Washington	2011	24 hours
United European Gastroenterology Week (UEGW), Amsterdam, the Netherlands	2012	16 hours
International Cancer Screening Network (ICSN), Rotterdam, the Netherlands	2015	16 hours
Seminars and symposia		
Research seminars at department of public health, Erasmus MC	2011-2016	100 hours
Nationaal symposium: Invoering van colonscreening, een scherpe blik vooruit, Zeist	2012	8 hours
Teaching		
Supervising medical students performing community projects as part of educational theme 3.C 'Arts en volksgezondheid', Erasmus MC, Rotterdam	2012-2015	100 hours

*1 ECTS = 28 hours


WITH SCREENING
Birth
Adult
Adult with adenoma
Stool test positive
Colonoscopy
Healthy senior
Death other cause
WITHOUT SCREENING
Birth
Adult
Adult with adenoma
Adult with carcinoma
Patient
Death due to CRC

List of publications

IN THIS THESIS:

Kapidzic A, van der Meulen MP, Hol L, van Roon AH, Looman CW, Lansdorp-Vogelaar I, van Ballegooijen M, van Vuuren AJ, Reijerink JC, van Leerdam ME, Kuipers EJ. Gender Differences in Fecal Immunochemical Test Performance for Early Detection of Colorectal Neoplasia. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015;13(8):1464-71.e4.

van der Meulen MP, Kapidzic A, Leerdam MEV, van der Steen A, Kuipers EJ, Spaander MCW, de Koning HJ, Hol L, Lansdorp-Vogelaar I. Do Men and Women Need to Be Screened Differently with Fecal Immunochemical Testing? A Cost-Effectiveness Analysis. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2017;26(8):1328-36.

van der Meulen MP, Lansdorp-Vogelaar I, Goede SL, Kuipers EJ, Dekker E, Stoker J, van Ballegooijen M. Colorectal Cancer: Cost-effectiveness of Colonoscopy versus CT Colonography Screening with Participation Rates and Costs. Radiology. 2018;287(3):901-11.

van der Meulen MP, Lansdorp-Vogelaar I, van Heijningen EM, Kuipers EJ, van Ballegooijen M. Nonbleeding adenomas: Evidence of systematic false-negative fecal immunochemical test results and their implications for screening effectiveness-A modeling study. Cancer. 2016;122(11):1680-8.

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, van Ballegooijen M, de Koning HJ, Lansdorp-Vogelaar I, van Leerdam ME; Socioeconomic differences in participation and diagnostic yield within the Dutch national colorectal screening programme with faecal immunochemical testing; *Submitted*

van der Meulen MP, Korfage IJ, van Heijningen EB, de Koning HJ, van Leerdam ME, Dekker E, Lansdorp-Vogelaar I; Interpretation and adherence to the updated risk-stratified guideline for colonoscopy surveillance after polypectomy - a nationwide survey.; *in Press*

OTHER PUBLICATIONS:

Arrospide A, Idigoras I, Mar J, de Koning H, van der Meulen M, Soto-Gordoa M, Martinez-Llorente JM, Portillo I, Arana-Arri E, Ibarrondo O, Lansdorp-Vogelaar I. Cost-effectiveness and budget impact analyses of a colorectal cancer screening programme in a high adenoma prevalence scenario using MISCAN-Colon microsimulation model. BMC cancer. 2018;18(1):464.

Idigoras I, Arrospide A, Portillo I, Arana-Arri E, Martinez-Indart L, Mar J, de Koning HJ, Lastra R, Soto-Gordoa M, van der Meulen M, Lansdorp-Vogelaar I. Evaluation of the colorectal cancer screening Programme in the Basque Country (Spain) and its effectiveness based on the Miscan-colon model. BMC public health. 2017;18(1):78.

Meester RG, Doubeni CA, Lansdorp-Vogelaar I, Jensen CD, van der Meulen MP, Levin TR, Quinn VP, Schottinger JE, Zauber AG, Corley DA, van Ballegooijen M. Variation in Adenoma Detection Rate and the Lifetime Benefits and Cost of Colorectal Cancer Screening: A Microsimulation Model. Jama. 2015;313(23):2349-58.

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. Quality Monitoring of a FIT-Based Colorectal Cancer Screening Program. Clinical chemistry. 2019;65(3):419-26.

Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I, 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. 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.

Penning C, Lansdorp-Vogelaar I, van Leerdam ME, van der Meulen MP, van Vuuren AJ, Kuipers EJ, Bonfrer JM, van Kemenade FJ, Biermann K, Thomeer MGK, Spaander VMCW, Kroep S, van Ballegooijen M, de Koning HJ; Landelijke monitoring van het Bevolkingsonderzoek Darmkanker: resultaten eerste half jaar 2014