Population-based colorectal cancer screening by fecal immunochemical testing over multiple rounds
van der Vlugt, M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary and future perspectives
SUMMARY AND FUTURE PERSPECTIVES

Colorectal cancer (CRC) is one of the major causes of death in the Netherlands, accounting for over 5100 deaths in 2015 [1]. Population screening aims to detect CRC or its precursors, advanced adenoma, in an early phase, thereby reducing CRC morbidity and mortality. Screening based on fecal occult blood testing (FOBT) using a guaiac FOBT (gFOBT) has been shown to reduce CRC mortality with approximately 15% [2], and a recent observational study demonstrated a 22% reduction in CRC mortality from screening with a quantitative and automated fecal immunochemical test (FIT) [3]. FOBT-based screening is a two-step screening strategy in which all individuals with a positive FOBT are referred for colonoscopy, the reference standard for the detection of CRC and adenomas [4]. FOBT-based screening can only be effective if there is a willingness to participate and perform the FOBT, as well as to subsequently undergo colonoscopy if FOBT-positive.

In the Netherlands, a national population screening program for CRC, based on biennial testing with a single FIT, is gradually being implemented since January 2014. Before the implementation of the national program, two FIT-based CRC screening pilot programs were initiated. Since 2006, these two pilot programs have been inviting approximately 27,000 persons for biennial FIT screening. In 2014, these two cohorts were combined and invited for a fourth pilot round of screening. Following this ‘forerunner group’ by continuing biennial FIT screening enabled us to obtain data over multiple screening rounds, and to anticipate the future results of the Dutch national program.

The research reported in this thesis covers a number of issues concerning four rounds of FIT-based CRC screening within this ‘forerunner group’ and aimed to evaluate the effectiveness of FIT-based CRC screening, in order to further optimize the screening strategy. We evaluated participation rates and adherence over multiple screening rounds, compared FIT-based screening to endoscopy screening, and studied cancers not detected within the screening program.

In the analysis presented in Chapter 1 we demonstrated that participation rates increased significantly over four rounds of FIT screening, from 60% to 63%, with significantly higher participation rates in women in all rounds. Of the 17,312 invitees eligible for at least two rounds of FIT screening, 72% participated at least once, whereas 28% never participated; 48% attended all rounds when eligible. Consistent participation was associated with older age, female sex, and higher socioeconomic status. We showed that uptake was high and increased after a reminder letter, not only after the initial invite, but also in subsequent screening rounds. We concluded that reminder letters should be sent to all non-responders in each screening round. However, almost a quarter of the invitees never participated in any of the rounds of FIT screening. It would be interesting to investigate how many of these invitees made an informed decision not to participate, and whether participation was hampered by barriers such as limited health literacy, distrust of government initiated health initiatives, cost considerations, or other issues.

Several CRC screening methods are available, varying in invasiveness and diagnostic accuracy. Of these screening modalities, fecal occult blood testing, with either the guaiac fecal occult blood test or the fecal immunochemical test, flexible sigmoidoscopy, and colonoscopy are used most often. Colonoscopy is considered the reference standard for detecting advanced neoplasia (AN): CRC and its advanced precursor lesions. As it is invasive, participation rates in primary colonoscopy
screening are generally low. FOBT-based screening and screening with flexible sigmoidoscopy both involve a two-step method: only if the first test is positive, a follow-up colonoscopy is advised. Whereas the diagnostic accuracy of the primary test is lower, an advantage of these programs is that participation rates are generally higher. In Chapter 2 we compared colonoscopy, flexible sigmoidoscopy and four rounds of FIT screening. We showed that participation was significantly higher with FIT screening (77%) than with flexible sigmoidoscopy (31%; P<0.001) and colonoscopy (24%; P<0.001). Among invitees, FIT detected significantly more AN compared to endoscopic screening: 4.5% (95% CI 4.2 to 4.9) versus 2.2% (95%CI: 1.8 to 2.6). Furthermore, FIT detected three times more CRC per invitee than endoscopic screening: 0.6% (95%CI 0.5 to 0.7) versus 0.2% (95%CI: 0.1 to 0.3). The number needed to scope to detect one AN was two for FIT, three for flexible sigmoidoscopy, and eleven for colonoscopy. Over multiple rounds, FIT detected more AN with the lowest number of colonoscopies, compared to endoscopic screening. We concluded that, in the Netherlands, FIT screening is the most effective screening strategy for CRC screening on a population level. As the results of this evaluation are highly driven by participation rates and these are known to vary between countries, it is possible that this conclusion does not apply to other countries.

When implementing a FIT-based CRC screening program, choosing the appropriate test can impact effectiveness. Several types of FIT are available on the market. All pilot studies until 2014 in the Netherlands have been performed with the OC-Sensor test (Eiken, Japan) whereas a tender selected a FIT from another company for the national Dutch screening program: the FOB-Gold (Sentinel, Italy). We performed a randomized comparison of the two tests in the fourth round of our pilot program. In Chapter 3, we showed that both the OC-Sensor and FOB-Gold test were returned by 63% of invitees (p=0.96). However, more non-analyzable tests were returned with the FOB-Gold, due to inappropriate use by the screenees. The positive predictive value was similar for the two tests, but the positivity rate was significantly higher for OC-Sensor, despite using the same cut-off. This led to a higher colonoscopy demand and a subsequently higher detection rate of advanced neoplasia for the OC-Sensor. When comparing both tests at the same positivity rate, instead of at the same cut-off, we observed similar predictive values and detection rates. Whereas comparisons between different types of FIT are usually done at the same Hb-concentration in the stool, our findings imply that the positivity rate rather than the Hb-concentration should be used for a fair evaluation of the diagnostic yield.

Although FIT screening can be expected to reduce CRC-related mortality, not all CRC will be detected within a screening program. In Chapter 4 we reported on screen-detected CRC and non-screen-detected CRCs in our pilot cohort. The proportion of interval cancers after three completed FIT-based screening rounds was 23%, reflecting a FIT sensitivity for detecting CRC of 77%. This proportion is much lower than the previously reported proportions of 48% to 55% in gFOBT-based screening. As expected, participants with FIT-detected CRC more often had early-stage tumors than those not detected within the screening program. Interestingly, survival of persons with a FIT interval cancer was comparable to the survival of persons with a screen-detected CRC, and better than that of non-participants and members of the general population with CRC. This implies that worries about falsely reassuring screening participants with a negative FIT result, though justified,
should not be overblown. One could hypothesize that these screening participants will not consult their GP, or postpone a visit, if they develop CRC symptoms in a later phase, thereby causing a delay in diagnosis and treatment, and potentially harming clinical outcome. Our results on staging and mortality in interval cancers do not suggest that this effect, if it exists, is large.

In over half of all participants undergoing colonoscopy after a positive FIT, no CRC or advanced adenomas are found. As fecal occult blood might also originate from the upper gastrointestinal (GI) tract, one could consider esophagogastroduodenoscopy to detect upper GI cancers in these persons with a negative colonoscopy after a positive FIT. In Chapter 5 we assessed the number of proximal cancers (i.e. oral cavity, throat, esophageal, gastric and small bowel cancer) within two years after FIT testing in screening: in FIT-positives, with and without advanced neoplasia at colonoscopy, and in FIT-negatives. After linkage, we identified 40 persons with a proximal cancer among 16,165 screening participants: 33 in FIT-negative participants (0.24%) and 7 in FIT-positive participants (0.33%). The difference was not statistically significant ($P = 0.43$). We concluded that esophagogastroduodenoscopy should not be routinely performed in FIT-positive screening participants without advanced neoplasia at subsequent colonoscopy.

As CRC screening programs are being implemented worldwide, an increasing number of early cancers (T1 cancers, with maximum depth of invasion into the submucosa) are being detected. These cancers should be recognized during colonoscopy, because they require a specific therapeutic approach. Several studies have shown that Asian experts can reliably recognize T1 cancers during colonoscopy. In daily practice, however, accurate endoscopic diagnosis of T1 cancers still seems challenging. In the study described in Chapter 6 we evaluated the performance of optical diagnosis of T1 cancers by European colonoscopy experts, general gastroenterologists and gastrointestinal fellows in an image-based study. Based on this image study both sensitivity for the optical diagnosis of a T1 cancer and negative predictive value for excluding a T1 cancer were found to be insufficient for all groups. Experts performed best, with a sensitivity of 67% and a negative predictive value of 78%. Our study indicates that training for endoscopic diagnosis for T1 cancers is urgently needed to ensure optimal clinical practice for treatment of these lesions.

**DISCUSSION AND FUTURE PERSPECTIVES**

Based on the studies reported in this thesis, with a focus on our ‘forerunner group’ of FIT-screening invitees, we could obtain data over multiple FIT-based CRC screening rounds. These data allow us to anticipate future results of the national screening program. In the Dutch national FIT-based CRC screening program, implemented since January 2014, all inhabitants aged 55 to 75 are invited for CRC screening by biennial FIT. The structure and logistics of the national program are mainly based on the results of our ‘forerunner group’.

The results of the studies reported in this thesis indicate that a FIT-based screening program, even based on sound pilot programs, should never be considered finished. In 2014, the first invitees of the national CRC screening program were invited for screening and, currently, those invitees are re-invited until the age of 75 for a second screening round. Linkage to the National Cancer Registry is being performed, but demonstrating a reduction in CRC-related mortality will take some time.
Since the start of the Dutch national screening program, a national information system has been developed for real-time monitoring all screening invitees. By closely monitoring the results of a screening program, adjustments can be applied quickly, if necessary, to optimize the effectiveness. There is always room for improvement, and studies should be initiated to further enhance the outcomes and effectiveness of screening, even when nationwide screening programs have already been implemented.

Within any screening program there will always be invitees that do not participate. This could be caused by limited health literacy, distrust of governmental health initiatives, because of costs, or due to other reasons. Previously performed questionnaire studies among screening invitees have shown that it is challenging to reach non-participants and therefore difficult to acquire information on their reasons for non-participation. It is conceivable that we need to diversify our invitation and information strategy, to achieve equity in the ability to make well-informed decisions about participation in CRC screening in all age groups and socioeconomic layers. One could consider providing additional instructive and informative videos via social media. We could also provide these videos in other languages, not only with the aim to reach non-native invitees but also functionally illiterate persons. Interactive platforms could be created, making it feasible for invitees to share questions raised, so these can be discussed and answered, increasing general knowledge about CRC and CRC screening.

As in every two-step screening method, a number of participants refuse colonoscopy after receiving a positive FIT-result, and some of them will eventually be diagnosed with CRC at a later point in time. Additional strategies to inform participants about the importance for follow-up after a positive FIT should be developed. Modifiable determinants for non-adherence to endoscopy, such as embarrassment, lack of knowledge about CRC, fear of the procedure, or inconvenience should be discussed with those who refuse colonoscopy. Additional strategies could be considered, e.g. sending an additional information leaflet to all non-responders after a positive FIT-result, or offering those persons a consultation by phone. A strategy involving general practitioners (GP) could also be an option. For example, the GP could be informed by the screening organization in case of non-adherence to colonoscopy after a positive FIT, and then contact the individual, to explore and discuss reasons for non-adherence.

As we showed, cancers will be missed in participants in FIT-based screening. Ideally, the sensitivity of FIT for CRC should be improved to further decrease the risk of missed lesions. However, this should not be at the cost of a significantly lower specificity, as this will consequently lead to an increased number of colonoscopies to be performed. Although the search for a superior, easy-to-use, in-expensive, non-invasive screening test is ongoing, at this time none of the current tests available yield the same sensitivity, specificity and ease of use as FIT. Stool DNA testing seems promising but further improvements in diagnostic accuracy and usability need to be performed before they can improve the efficacy of CRC screening programs.

Apart from striving to maintain high participation rates over consecutive FIT-screening rounds, another method to increase the sensitivity of FIT-based screening is to combine FIT-screening with risk stratification. This implies the identification of individuals at increased risk of having advanced neoplasia by using a set of risk factors, which, when combined with the quantitative FIT-result,
SUMMARY AND FUTURE PERSPECTIVES

leads to a probability of having advanced neoplasia. Inviting screening participants at increased risk, rather than FIT-positives, may improve the detection of advanced neoplasia in FIT-based CRC screening without increasing the number of colonoscopies, thereby reducing the number of negative colonoscopies. Known risk factors for advanced neoplasia are high age, male sex, a positive family history for CRC, and active smoking. Collecting this information from participants would be simple, in principle, but should not jeopardize invitees’ willingness to participate. A study protocol to evaluate this risk-based variant of FIT screening has been developed and this study will hopefully start shortly. It will be performed within the setting of our national FIT screening program.

In the future, even further tailored personalized screening could be implemented, based on more individual risk factors and features, such as fecal Hb concentration found in previous rounds, or screening behavior of the participant (consistent vs. inconsistent). Tailoring screening intervals is another risk-based approach that could be used.

Post-colonoscopy cancers are strongly operator-dependent, reflecting the quality of the colonoscopy. Most of these should be preventable. Studies have reported interval cancer rates varying between 2% and 8% [5-9]. Lesions could be missed, for instance due to inadequate bowel preparation or incomplete intubation. At the start of the Dutch CRC screening program, large efforts have been made to assure colonoscopy quality as delivered to FIT-positive individuals, who carry a high probability of advanced neoplasia. Based on scientific data on quality parameters, a formal accreditation process and regular audits for centers, as well as for individual endoscopists, were developed and implemented. With this strategy, we hope to ensure lower rates of post-colonoscopy CRCs within the national CRC screening program than observed in our pilot cohort.

To preserve and enhance colonoscopy quality even more, polypectomy techniques could be further improved and standardized, to lower the risk of incomplete removal as a cause for interval cancers. Furthermore, technical improvements to the endoscopes could improve the ability to assess radicality after polypectomy and facilitate optical diagnosis. As the incidence of T1 CRCs increases, because of the widespread implementation of CRC screening programs, adequate endoscopic recognition of early CRCs is of importance. These cancers can be cured by endoscopic treatment only if polypectomy is performed appropriately. Endoscopists should increase their skills in optical diagnosis to recognize early cancers and treat those appropriately. This could be facilitated by offering a formal training in optical diagnosis and all (new) suitable treatment options for early CRCs.

Several questions concerning consecutive rounds of FIT-screening have been answered in this thesis, but many others still need further investigation. With ongoing research within the existing screening program we hope to optimize the basic promise to screening invitees: a reduction in colorectal cancer morbidity and mortality from participation in FIT-based screening.
REFERENCES

1. Dutch Cancer Registry. www.cijfersoverkanker.nl