Chapter 8

Summary, general discussion and future perspectives.
SUMMARY AND GENERAL DISCUSSION

Colorectal cancer (CRC) screening by guaiac based faecal occult blood tests (g-FOBTs) was shown to decrease CRC related death.\(^1\)\(^-\)\(^3\) In this screening setting, subjects with a positive test were referred for colonoscopy. Later, test performance of faecal immunochemical tests (FITs) was shown to be superior over g-FOBT.\(^4\)\(^-\)\(^7\) This thesis explored different aspects of FIT performance that influence CRC screening and that were previously unknown. The studies presented were performed in outpatients referred for colonoscopy. Therefore, several research questions were answered that can hardly be answered in a screening setting. The results are complementary to the current knowledge on FITs for CRC screening.

The study presented in Chapter 2 compared FIT results in CRC cases from two studies with a different population and study design. It showed that a FIT result is largely dependent on tissue tumour stage. The higher FIT results in the referral population are a reflection of a higher prevalence of advanced stage CRC. Pilot screening programmes may provide the best context to study screening tests, since the tests are evaluated in the population intended to be screened in the future. However, due to its design, some research questions cannot be answered. The study presented in Chapter 2 indicates that after correction for tissue tumour stage distribution, both screening and referral cohorts provide similar test results. This supports the use of elective or clinical settings to test hypotheses in an early phase of research, when a large number of outcome variables is desired, or when colonoscopy is required in all participants. Although this approach is not a substitute for screening studies, at substantially lower costs and in less time, relevant research questions can be answered in referral populations. In addition, by merging results from different sources, the strength of the evidence can be enlarged.

As described in previous studies, FIT sensitivity can be increased by performing multiple tests.\(^8\)\(^,\)\(^9\) This is not surprising, as any additional selection method that will result in more colonoscopies will lead to a higher detection level of advanced neoplasia. However, as described in Chapter 3, the increase in sensitivity is counterbalanced by a lower specificity. When sampling two FITs was compared with sampling one FIT at a fixed value for specificity, a comparable sensitivity was found at a different cut-off level when using a single test. This seems to be in line with an Italian study where detection rates increased with two FITs compared to one FIT, and positive predictive values decreased accordingly.\(^8\) Recently, our findings were confirmed by French investigators who compared sampling one FIT with
two FITs in a screening setting. At a different cut-off value, the ratio of sensitivities and false positive tests (as surrogate endpoints for sensitivity and specificity) were found to be similar.\(^\text{10}\) In addition, a study simulating screening with one and two FITs found that sampling one FIT is likely to be more cost-effective than double sampling.\(^\text{11}\) Therefore, when a FIT screening programme requires a higher sensitivity, the preferred solution is to use one test at a lower cut-off value.

It is known that haemorrhoids can cause rectal bleeding and their prevalence may be high, as one report found a prevalence of 86%.\(^\text{12}\) Therefore, haemorrhoids could lead to many false positive tests. If this problem indeed exits one might consider a different screening test or cut-off value for subjects known to have haemorrhoids. However, in Chapter 4 it was shown that the presence of haemorrhoids in subjects performing a FIT was not associated with a higher rate of false positive test results. Subjects with haemorrhoids did not have more false positive results than subjects without any abnormalities in the colon. As average risk subjects are likely to have a lower prevalence and less severity of haemorrhoids compared with the high risk individuals studied, the present study may represent the worst case scenario. The actual frequency of false positive tests due to haemorrhoids may therefore be even lower.

The sensitivity of FIT for CRC is higher and the specificity is lower in males compared with females, and thereby the results presented in Chapter 5 are in agreement with previous work.\(^\text{13}\) This difference was not caused by known confounders as age and location of CRC in the colon. The same sensitivity for CRC as in males could be reached by lowering the cut-off value in females. However, whether sex specific cut-off values are needed is questionable. Females may have a lower sensitivity of FIT but have a higher participation rate, as was found in the English and Scottish screening programmes.\(^\text{14, 15}\) In addition, prevalent cancers will presumably be detected in the first screening rounds, and the focus of screening may shift to detection of advanced adenomas. As only a small gender difference in the sensitivity for advanced adenomas was found, gender specific cut-off values seem unnecessary. The actual benefit of tailored screening depends on several factors and modelling studies may provide insight in this matter.\(^\text{16}\) It should be kept in mind that individualising screening guidelines adds to the complexity of a screening programme and should only be adopted if the expected benefits are substantial.

Under the assumption that the CRC screening programme in the Netherlands would be using OC Sensor\(^\text{®}\) as primary screening test, the total colonoscopy burden after complete roll-out of the programme was calculated. It
was found that the current 191,339 colonoscopies per year need to be increased to a total of 269,339 colonoscopies yearly (an additional 41%). The study described in Chapter 6 did not investigate rest capacity, but the extra workload will anyhow increase pressure on current capacity. Assuming that the colonoscopies will be performed in the province of residence, 2 to 4.5 extra colonoscopies need to be performed daily. Due to the stepwise implementation of the programme, there is time to create more capacity by e.g. training more endoscopists, supporting staff and increasing means.

Another way to increase capacity could be by adjusting surveillance programmes, as 30% of colonoscopies in the Netherlands are performed because of surveillance for adenomas, CRC or a positive family history. In Chapter 7 it was explored whether using FIT as pre-selection before colonoscopy in surveillance is accurate enough to exclude presence of advanced neoplasia. Unfortunately, this seems an unfeasible option, as 70% of advanced neoplasia were missed. However, future studies could focus on (epi)genetic alterations involved in adenoma to carcinoma progression in adenomas detected in surveillance. This could result in a personalised risk classification that may guide surveillance intervals. In addition, when molecular biomarkers could detect the genetic alterations in stool from subjects under surveillance, a stool test could be used to guide surveillance intervals. Such a test might result in less adenomas missed, and due to a potential higher specificity a decrease in the strain on capacity.

Screening is a realistic approach to decrease mortality from CRC. In the Netherlands, screening with the gold standard, colonoscopy, is no longer considered to be an option because of the low participation, the risk of complications, the burden of the procedure, the costs, and the knowledge that 95% of all inhabitants will never get CRC. It is therefore of vital importance to have a valid, stable and accurate pre-selection test to identify those subjects at increased risk for presence of precursor lesions or CRC. FIT is easy to perform at home, cheap, non invasive, and is known to have a relatively high participation rate. Although it has moderate test characteristics, biennial repetition of the test is very likely to result in mortality reduction and potentially also in a decrease in incidence of CRC. The start of the Dutch national CRC screening programme was received with a lot of enthusiasm, but screening accuracy should be improved substantially. What should future research aim at, and how can we take CRC screening to the next level?
FUTURE PERSPECTIVES

Although nationwide CRC screening has started and CRC related mortality is expected to decrease by 2,400 cases each year after full roll-out over the Netherlands, improvement of the programme is desirable from the start.

Firstly, it should be clarified which FIT is the best to use in the screening programme. The tendering programme in which the FIT for screening was selected, yielded the FOB Gold®. This is a different test than the FIT for which most evidence of effectiveness is available; OC Sensor®. Comparative studies, especially of high quality, on both tests are sparse. Recently, a comparative study from Spain showed a 3% higher participation rate, a 2% lower positivity rate, but a higher number of true positive test results for OC Sensor®. The frequency of error in performance was higher for FOB Gold®. However, both tests were not compared at different cut-off values, neither at an equal predetermined amount of haemoglobin in the test. As it is known that FITs from different manufacturers differ in test characteristics, it is preferred to perform a study directly comparing at least FOB Gold® and OC Sensor® at a cut-off value with an equal amount of haemoglobin per millilitre, as well as at different cut-off values. Even small differences may have major impact in a screening population of millions of people. As shown by this thesis, outpatients referred for colonoscopy can be invited for such a diagnostic comparison study, as tests can be compared at relative low costs and in little time, and test accuracy can be determined in a high number of carcinomas and advanced adenomas.

Secondly, the currently available FITs seem to have reached their maximum performance, but need to be improved further. This maximum performance is far from perfect. Faecal immunochemical tests have a low sensitivity for right sided lesions, have moderate overall test characteristics (especially for advanced adenomas), and are non-specific in that they detect human blood, not high risk neoplasia. Detection levels can be increased by lowering the cut-off value or by using multiple tests, but always at cost of a lower specificity. The detection rate could theoretically be further increased by using anticoagulants before FIT sampling. Hereby, the chance that advanced neoplasia will bleed might increase. Another option could be to repeatedly perform FIT in a short interval (e.g. once every month for six months), in an attempt to correct for intermittent blood loss. However, it is unknown whether all neoplasia bleed, whether continuously or intermittently. In addition, higher detection levels of blood will not necessarily result in a higher efficiency of screening, as there are multiple causes of occult bleeding.
Efficiency of CRC screening could be improved in two ways: 1. improvement of test characteristics by detection of tumour specific products in addition to bleeding; 2. detection of those colonic neoplasia that actually carry high risk of progression to carcinoma. It is known that carcinomas, as well as adenomas, shed and exfoliate cells and cell substances as DNA into the bowel lumen. These waste products can subsequently be detected in stool. This material could be the fingerprint of adenomas developing to carcinoma, and provide the opportunity to detect non-bleeding lesions. Tumour specific markers in combination with FIT (FITplus) might close the gap left by FIT. A pilot study showed the potential of such a combination test; test sensitivity increased and specificity was hardly affected. Secondly, efficiency of screening could be increased by using tumour specific stool markers that detect only those adenomas developing towards carcinoma. By detection of (epi)genetic changes known to occur in the adenoma to carcinoma sequence, overdiagnosis and overtreatment of harmless adenomas could be decreased. Screening could become more efficient when development of a new generation of biomarkers could accurately detect true high risk lesions. A shift in the paradigm of what we call advanced or high risk adenomas is needed. By a change towards the biology of carcinogenesis, the accuracy and efficiency of screening can further be improved.

In the Netherlands, CRC screening started in 2014 by using FOB Gold®. As the Dutch Health Council recommends, it should be possible to compare the operational programme to new or different screening tests, as new developments may become more efficient. The addition of promising tumour specific biomarkers such as PHACTR3 would be an excellent opportunity for improvement. After a few rounds of screening it is likely that prevalent carcinomas have been detected, and the aim of screening may shift towards detection of high risk adenomas. The paradigm should shift to detect only adenomas at high risk to carcinoma progression. By making FIT more targeted to carcinogenesis, we take it to the next level.
REFERENCES


