1 General introduction
This thesis focuses on screening strategies to reduce the burden of colorectal cancer (CRC) by means of early detection of both tumors and CRC precursor lesions. Using mathematical modelling, we evaluate the impact of screening on CRC incidence and mortality. Besides estimating the positive health effects of screening, we assess resource use associated with screening and screening burden.

In this first Chapter, we will introduce this thesis by briefly describing the epidemiology of CRC and the development of tumors in the colon and rectum. Subsequently, an overview of available screening tests will be presented as well as current insights concerning the long-term effectiveness of screening. Furthermore, the value of mathematical modeling in the implementation and evaluation of CRC screening programs will be discussed. We will end with an outline of the following Chapters of this thesis.

COLORECTAL CANCER

Epidemiology

Colorectal cancer (CRC) affected roughly 1.4 million individuals worldwide in 2012, making it the third most common cancer in men and the second most common cancer in women. Although there is wide geographical variation, CRC is more common in affluent countries. With approximately 694,000 deaths in 2012, CRC accounted for 8% of all cancer deaths. Mortality rates are slightly higher in low-income countries, indicating lower survival probabilities. Both incidence and mortality rates are higher for men than for women. Also older individuals have an increased risk of CRC.

Also in the Netherlands, CRC is an important health problem. In 2013, there were over 13,300 new cases and almost 5,000 individuals died due to this disease. Without intervention, CRC burden is likely to rise in the coming years. This is partly due to population aging, because the risk of CRC is age-dependent. Population aging, together with population growth, is expected to increase CRC incidence with over 40% in the coming twenty years. In addition, CRC risk factors such as physical inactivity, red and processed meat consumption, low intake of fruits, vegetables and fiber and overweight and obesity are becoming more prevalent.

Natural history

CRC arises from precursor lesions located in the bowel. The development from a precursor lesion to cancer is a multi-step process, which is driven by genetic alterations. There are two types of precursor lesions; adenomas and serrated lesions. The development of an adenoma to CRC is described as the adenoma-carcinoma pathway whereas the progression from serrated lesion to cancer is defined as the serrated pathway.

Adenoma-carcinoma pathway

The adenoma-carcinoma pathway was postulated by Morson in 1974. According to this pathway, adenomas, which are benign tumors of glandular origin, can transform into malignancy. A subgroup of adenomas, i.e. advanced adenomas, is thought to have increased malignant potential. Advanced adenomas are defined as adenomas that are ≥ 1 cm, have villous architecture and/or have high-grade dysplasia. Adenomas can be pedunculated (with a stalk), sessile (broad-based) or flat (not or only slightly elevated). However, the latter are uncommon. Adenomas can grow in size, but regression is also possible. Only few studies have assessed the growth rate because in general, adenomas are removed upon detection. Limited evidence suggests that diminutive adenomas, i.e. <6 mm, have
a net tendency to grow in size whereas small adenomas, i.e. 6-9 mm, have a net tendency to regress in size. Furthermore, advanced adenomas may grow more rapidly than non-advanced adenomas. Results regarding complete regression of adenomas are inconclusive. Two endoscopy studies with two and three years of follow-up reported no complete regression whereas two more recent computer tomographic colonography (CTC) studies with a similar follow-up period did observe spontaneous regression of adenomas. However, it is possible that adenomas were overlooked in the latter studies or that the initial observed adenoma was the result of a false-positive CT-scan. Besides changes in size, an adenoma may acquire villous characteristics and/or high-grade dysplasia over time. The risk of developing such histological features increases with adenoma size.

The duration of the development from an adenoma into CRC, defined as the adenoma dwell time, is thought to take several years. However, the exact adenoma dwell time is unknown. Since the prevalence of adenomas is markedly higher than the prevalence of CRC, only few adenomas develop into cancer. Limited evidence suggests that only 5% of adenomas transition into malignancy.

### Serrated pathway

The malignant potential of serrated lesions is only recently recognized. As a consequence, there is more uncertainty regarding the development of serrated lesions to CRC compared to the extensively studied adenoma-carcinoma pathway. Serrated lesions are divided into three subtypes, namely hyperplastic polyps (HPs), sessile-serrated adenomas (SSAs) and the extremely rare traditional serrated adenomas (TSAs). The relationship between HPs and SSAs is unclear; it is hypothesized that HPs and SSAs may arise independently whereas it is also suggested that HPs develop into SSAs. The former hypothesis is based on differences in location; the majority of HPs are located in the distal colon whereas SSAs are often located in the proximal colon. The latter theory is underpinned by histological similarities of HPs and SSAs, i.e. a saw-toothed appearance. Only SSAs are thought to progress to CRC but the proportion of CRCs that arises via the serrated pathway is unknown. Estimates based on genetic alterations detected in both CRCs and serrated lesions vary between 5% and 30%. Accumulating evidence guides these estimates towards 15% or more of CRCs.

### PREVENTION AND EARLY DETECTION OF COLORECTAL CANCER

The benign precursor lesions and the long preclinical phase make CRC an excellent target for screening. The benefit of CRC screening is two-fold. Firstly, early detection and removal of precursor lesions disrupts the pathway to CRC, thereby preventing the development of cancer. Secondly, screening can increase survival probabilities due to detection of tumors in an earlier stage.

### Screening tests

Several screening modalities are available for the detection of colorectal lesions and cancer. The test that is considered the gold standard is colonoscopy. During this procedure, the entire bowel is visually inspected by a camera on a flexible tube that is passed through the rectum. If colorectal lesions are detected, these can be removed by means of polypectomy. The detection rate for adenomas is high; respectively 74%, 87% and 98% of diminutive, small and large adenomas are identified. Detection rates for serrated lesions have not been reported but these are likely to be lower compared to those for adenomas because their pale color, proximal location and flat appearance hamper visual detection. Tumors are rarely missed by colonoscopy. Due to the high detection rates, the interval between subsequent screening colonoscopies may be as long as ten
years. As a consequence, it is sufficient to offer few screening colonoscopies in the lifetime of an individual.

The meta-analysis by Khalid-de Bakker et al. (2011) reported that first-time participation rates for screening colonoscopy vary between 16% and 37%, with an average first-time participation rate of 26%. This wide range is caused by the fact that screening participation is highly dependent on invitational methods, public awareness campaigns and year in which individuals were invited for screening. In the Netherlands, a colonoscopy screening trial reported a participation rate of 22%. In general, participation rates for colonoscopy are rather low due to the invasiveness of the procedure. Furthermore, the bowel preparation is considered burdensome.

Stool-based tests such as the guaiac fecal occult blood test (gFOBT) and the fecal immunochemical test (FIT) are less invasive. gFOBT and FIT are assays which detect respectively haem and human globin in stool that originates from colorectal lesions and tumors. FIT has several advantages over gFOBT. Firstly, FIT requires no dietary restrictions since globin is human-specific. Thus in contrast to gFOBT, traces of dietary blood will not lead to a positive test outcome. In addition, one sample is sufficient whereas multiple samples need to be provided for gFOBT. Lastly, FIT has a higher sensitivity for advanced adenomas and CRC. Still, FIT misses around 20% of cancers and 70% of advanced adenomas. For serrated lesions, the limited evidence available suggests that they are less likely to bleed compared to adenomas. Thus, FIT may also miss the majority of these lesions.

The probability of detecting a cancer or a relevant precursor lesions increases by repeated testing. The low sensitivity of stool-based tests necessitates a short screening interval. In general, the interval is set at two years leading to a high number of screens offered in the lifetime of an individual. A positive stool-based test is followed by referral to diagnostic colonoscopy.

Due to the non-invasiveness of stool-based testing, participation is considerably higher than for colonoscopy screening. First-time participation rates reported by Khalid-de Bakker et al. were 51% (range 15%-81%) and 43% (range 17%-62%) for gFOBT and FIT, respectively. In the Netherlands, trials assessing acceptance of stool-based tests reported that participation rates were around 50% for gFOBT whereas participation rates for FIT were around 60%.

Other screening tests are sigmoidoscopy and computed tomographic colonography (CTC). Sigmoidoscopy is a similar technique to colonoscopy, but only the distal part of the bowel is inspected. With an average of 52% (range 14%-63%), participation is considerably higher than for colonoscopy. CTC is an imaging technique and enables external inspection of the colon. Consequently, CTC is less invasive than colonoscopy. In addition, a more limited bowel preparation is required. Based on these two factors, it would be expected that participation rates for CTC are higher than for colonoscopy. Nevertheless, Khalid-de Bakker et al. reported an overall participation rate of 22% (range 16%-24%), i.e. comparable to colonoscopy. However, some of the studies on CTC included in this meta-analysis used full bowel preparation, possibly leading to lower participation rates for CTC. A recent Dutch trial comparing colonoscopy and CTC found a participation rate of 34% for CTC with limited bowel preparation, as compared to 22% for colonoscopy. Furthermore, participants in this study favored CTC over colonoscopy.

New screening techniques such as DNA testing and colon-capsule endoscopy are evolving. DNA tests detect genetic mutations that are associated with CRC and its precursor lesions in either blood or feces. While blood-based DNA testing is still under development, stool-based DNA testing is already approved by the U.S. Food and Drug Administration. A study comparing a stool-based DNA test with FIT in average risk individuals reported that the DNA test had higher sensitivity for cancers
and advanced adenomas than FIT. However, the specificity of the DNA test was 94.9% versus 96.4% for FIT, leading to more false-positive test results. In the context of screening, a slightly lower specificity causes considerable burden because most individuals do not have adenomas or CRC.

With colon-capsule endoscopy, individuals swallow a small capsule with a camera that moves through the bowel by peristalsis. To apply this method, good colonic preparation is imperative because the bowel cannot be cleansed during the procedure. For both colon-capsule endoscopy and DNA testing, additional research on diagnostic accuracy and participation rates is required.

**Long-term screening trials**

Several trials have evaluated the long-term effects of gFOBT screening on CRC mortality. Six to seven rounds of biennial gFOBT reduced CRC mortality by 15% after eleven to thirteen years of follow-up. When follow-up was twenty years, three to eleven rounds of biennial gFOBT screening led to a 13% reduction in CRC mortality. The Minnesota trial, consisting of six rounds of biennial gFOBT, had the longest follow-up period, i.e. thirty years. In this study, CRC mortality was 22% lower in the screened group than in the control group.

The long-term effects of sigmoidoscopy screening have been evaluated as well. A Norwegian study reported that once-only sigmoidoscopy led to a 28% and 16% reduction in CRC incidence and mortality, respectively, after eleven years of follow-up. Two other studies concerning once-only sigmoidoscopy with a similar follow-up period showed slightly lower reductions for CRC incidence, i.e. 18% and 23%, whereas mortality reductions were higher with 22% and 31%. In a trial with eleven rounds of follow-up, two rounds of sigmoidoscopy screening reduced CRC incidence by 21% and mortality by 26%.

Data on the impact of screening colonoscopy on CRC incidence and mortality is not yet available, results of the first trials are expected within several years from now. Nevertheless, prospective cohort studies have shown a considerable reduction in CRC mortality after more than twenty years of follow-up.

Also for FIT screening, the impact on CRC incidence and mortality has not yet been assessed since FIT is a relatively new test. Currently, three studies are ongoing concerning multiple rounds of FIT screening. However, follow-up is still too short for conclusions on screening impact. It is expected that FIT screening will lead to larger incidence and mortality reductions than gFOBT due to the higher sensitivity for CRC and advanced adenomas.

**National screening programs**

Since screening is an effective tool to decrease CRC burden, various countries have implemented CRC screening. Some countries have opted for opportunistic screening, i.e. only patients who are visiting the GP on their own initiative are invited for screening (e.g. Germany, Austria and Switzerland). Other countries have chosen for an organized screening program (e.g. the Netherlands, Finland and the UK). Screening programs vary widely in screening test used, screening age range and screening interval. These differences are due to, among others, variations in CRC incidence, the range of available screening tests, available health care resources and health care organisation.

The Netherlands have recently started with the implementation of an organized screening program consisting of biennial FIT screening in individuals aged 55 to 75 years. Individuals with a positive test are referred to diagnostic colonoscopy during which detected lesions are removed. Based on the findings at colonoscopy, individuals are allocated to a low, intermediate and high risk
group. Individuals in the low risk group return to the screening program after ten years. Individuals in the intermediate and high risk group enter the surveillance program and are offered a surveillance colonoscopy after respectively five and three years.65

COLORECTAL CANCER SCREENING MODELS

The choice for a CRC screening program, i.e. screening test, screening age range and screening interval, can be difficult. The optimal screening program in terms of benefits and costs can be determined using mathematical models. These models extrapolate short-term effects on intermediate outcomes, as observed in clinical trials, to long-term CRC incidence and mortality. With these models, numerous screening strategies can be compared.

Model structure

There are quite a number of models that can be used to evaluate CRC screening strategies. Some of these models are relatively simple decision trees, as for example described by Joseph et al. (1988).66 The majority of the models in use, however, are natural disease models, describing the progression of adenomas to CRC via the adenoma-carcinoma pathway.67–76 This pathway can either be characterized as a stepwise progression from “healthy” to “low-risk adenoma” to “high-risk adenoma” to “CRC”,67,73,76 or as growth in size, i.e. transition from “healthy” to “small adenoma” to “large adenoma” to “CRC”.68,69,75 In the former representation, size is not taken into account. Since test characteristics of screening tests are size-dependent, this could lead to inaccurate model predictions. Only few models include histological features of adenomas such as dysplasia and villosity,67,75 which may indicate adenomas with increased malignant potential, thereby representing the natural history of CRC more accurately.

By including only the adenoma-carcinoma pathway to CRC, existing models assume that all CRCs arise from adenomas. Thus, the serrated pathway to CRC is not taken into account. This is due to the fact that serrated lesions were thought to be innocuous at the time most of these models were developed.28 However, three of the published models include the possibility of ‘de novo’ cancers, assuming that a limited proportion of CRCs arises without a (known) precursor lesion.70,75,76 The idea of ‘de novo’ cancers is possibly based on an unexpectedly large proportion of interval cancers in individuals screened with colonoscopy.

Parameterization

Mathematical models provide a framework in which all available evidence on CRC can be synthesized. Some model parameters can be derived directly from studies, such as the location distribution of colorectal lesions and the proportion of adenomas with certain morphology. Other model parameters cannot be observed directly. An example of such a parameter is the adenoma dwell time; since adenomas are removed upon detection, this parameter cannot be observed in practice. The uncertainty regarding this parameter is reflected by the differences in adenoma dwell time between models; the dwell time of three well-known models varies between 8 and 24 years.77

Previous research has shown that adenoma dwell time is an important parameter; it has substantial impact on predicted screening effectiveness.77 Models with a short adenoma dwell time will favor strategies with a short screening interval because a long screening interval will allow precursor lesions to develop into CRC. In contrast, models with a long adenoma dwell time will favor strategies with a long screening interval, if the test is sufficiently sensitive, because less
(costly) screening rounds are required while still detecting the majority of adenomas before they progress to CRC.

The adenoma growth rates in most CRC screening models, a parameter that partly determines the adenoma dwell time, were calibrated to adenoma prevalence data from autopsy studies.\textsuperscript{78–87} These studies are dated; the majority was conducted more than 25 years ago. Due to increasing exposure to CRC risk factors in the last couple of years, current adenoma prevalence is likely to be higher. In addition, the possibility of bias due to selective samples should be acknowledged.\textsuperscript{88} Furthermore, these studies were not always conducted in the countries for which the model is used. Consequently, the estimated incidence and growth rates are likely to deviate from the real adenoma incidence and growth rates, leading to biased model outcomes.

**Effectiveness and cost-effectiveness analyses**

Model-based predictions of the he long-term effects of screening on CRC incidence and mortality have been reported in two studies.\textsuperscript{89,90} The Dutch study predicted that thirty years of biennial FIT screening in individuals aged 55 to 75 years may prevent 2,270 CRC-related deaths per year.\textsuperscript{89} For Germany, thirty years of primary colonoscopy screening in individuals aged 55 years and older was predicted to prevent over 14,000 CRC cases per year.\textsuperscript{90} The impact of the German screening program will keep increasing until forty years after the implementation. From that point onwards, around 16,000 CRC cases are predicted to be prevented each year.

Many modeling studies have evaluated the cost-effectiveness of CRC screening. Screening cost-effectiveness studies compare health and cost consequences of two or more screening strategies.\textsuperscript{91} In the evaluation of CRC screening, health consequences are often expressed as life years lived whereas costs refer to the monetary value of the resource use due to a specific screening strategy. Subsequently, a strategy is compared to the reference strategy by calculating the incremental cost-effectiveness ratio (ICER). That is, the difference in costs is divided by the difference in live years lived. When a strategy leads to more live years lived than the reference strategy and the ICER is under a pre-specified cost-effectiveness threshold, the strategy is considered cost-effective compared to the reference strategy.

An overview of the cost-effectiveness studies on CRC screening is provided by Lansdorp-Vogelaar et al. (2011).\textsuperscript{92} All studies included in this overview reported that CRC screening is cost-effective or even dominant, that is more effective and less costly, compared to no screening. However, there is no consensus regarding the optimal screening modality. On the other hand, this review clearly showed that emerging screening tests such as stool-based DNA testing and colon-capsule endoscopy are not yet cost-effective compared to more established screening methods.

**AIM OF THIS THESIS**

The aim of this thesis is to predict the long-term impact of the Dutch FIT-based CRC screening program and assess possible threats. Furthermore, we explore opportunities to optimize such a screening program.

**Outline of this thesis**

In Chapter 2, we describe the development of a mathematical model that includes both the adenoma-carcinoma pathway as well as the serrated pathway to CRC. This model allows evaluation of the impact of the serrated pathway on predicted CRC screening effectiveness. An important data source for the model was a Dutch screening trial in which asymptomatic, never-screened individuals
aged 50 to 75 years were subjected to colonoscopy.\textsuperscript{33} The model was named the Adenoma and Serrated pathway to Colorectal CAncer (ASCCA) model.

In Chapter 3, we study the impact of varying assumptions concerning the serrated pathway on the long-term effectiveness of the Dutch CRC screening program using the ASCCA model. Among other things, we set the contribution of the serrated pathway to CRC incidence to 0\%, 15\% and 30\% and predicted, for each scenario, the CRC incidence and mortality after thirty years of screening. These predictions were compared to a no screening scenario.

The Dutch CRC screening program consists of eleven rounds of FIT screening. This high number of screening rounds is required because of the low sensitivity of FIT for relevant CRC precursor lesions.\textsuperscript{38–41} However, such a large number of screening rounds may lead to screening fatigue. That is, individuals may lose the motivation to participate in screening because they already participated several times, which may lead to decreased participation rates. We study the potential impact of screening fatigue on overall program effectiveness in Chapter 4.

In Chapter 5 and 6, we explore opportunities to optimize the Dutch CRC screening program. In Chapter 5, we compare several screening tests in terms of effectiveness, costs and cost-effectiveness. We focus on imaging techniques because they have comparable diagnostic accuracy as colonoscopy but are less invasive and require only limited bowel preparation.

In Chapter 6, we study the added benefit of surveillance in the context of an implemented screening program. Evidence concerning the optimal surveillance strategy to reduce future CRC risk is limited. To gain more insight, we compared several screening plus surveillance strategies to a screening without colonoscopy surveillance strategy.

Chapter 7 summarizes the main findings presented in this thesis. In addition, we discuss methodological issues and provide recommendations for future research.
REFERENCES

49. Spada C. Colon capsule endoscopy: What we know and what we would like to know. World J Gastroenterol 2014; 20: 16948.