Optimizing Outcomes of Colorectal Cancer Screening

Reinier G. S. Meester

Optimizing Outcomes of Colorectal Cancer Screening

Reinier G. S. Meester

ISBN: 978-94-6169-997-8
Printing and layout: Optima Grafische Communicatie, Rotterdam, The Netherlands

Optimizing Outcomes of Colorectal Cancer Screening

Het optimaliseren van de uitkomsten van darmkankerscreening

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnficus
Prof. dr. H.A.P. Pols
and in accordance with the decision by the Doctorate Board.
The public defense shall be held on
Wednesday 4 January 2017 at 11.30 am

by

Reinier Godefridus Steven Meester
born in Hoorn

Erasmus University Rotterdam

DOCTORAL COMMITTEE

Promotor:	Prof.dr. H.J. de Koning
Other members:	Prof.dr. M.J. Bruno
	Prof.dr. C.A. Uyl-de Groot
	Prof.dr. A.A.M. Masclee
Copromotor:	Dr. I. Lansdorp-Vogelaar

FOREWORD

Rotterdam, Netherlands, October 2016.

Before you is what can be viewed as the culmination of four years of dedicated work. Four years ago I learned about a vacancy at the Erasmus Medical Center and met my current co-promotors. Not much later I decided to make the unorthodox transition from abstract financial predictions to public health research. This meant also a transition from business to academia, from clear targets and deadlines to a greater sense of independency, from national to international collaborations, from Amsterdam to Rotterdam. There has not been a single day that I have regretted the decision.

For the few who will actually read through some of the remainder of this thesis, I hope that you will find some interesting new ideas and methods to inform your research or practice. The thesis opens with a general introduction into the field of public health and the subject of colorectal cancer screening in particular, followed by three parts in which the main research findings are presented, and is concluded by a general discussion of the most important findings and implications. All chapters in the three middle parts can be read independently in combination with Chapter 2. I have tried to be consistent throughout in the use of terminology and abbreviations, and have tried to remove repetitious text on methods and background. In some instances, however, this was not possible without reformulating the original study reports as published in the literature, in which case I favored to preserve the original text. I hope and expect that this will not cause any serious confusion.

I am humbled to have been able to collaborate with many excellent researchers over the years both within our department as well as abroad. I have learned a lot from these collaborations, and owe a great deal to my co-authors. Although I will make acknowledgements at the end of this thesis, I would like to briefly mention two names here who were particularly important for this thesis, Dr. Marjolein van Ballegooijen and Dr. Iris Lansdorp-Vogelaar. Thank you for the trust and opportunities that you have given me from the start. I could not have hoped for a better first step into this field of research.

To the question whether the thesis will really add significantly to the knowledge that the Universe contains, I would echo Dr. Francis Collins, director of the US National Institutes of Health in Nature: “Well, it would be a rather small contribution, to be sure. I think the greatest beneficiary of my PhD was not the Universe, it was probably me.”

CONTENTS

Part I

Background

Chapter 1

General introduction

Colorectal cancer is a first-order global public health problem. It is the fourth leading cause of cancer-related deaths among men and women,[1] despite an increasing awareness among researchers, policymakers and public. The disease is associated with Western diet, and primarily affects Western countries (**Figure 1.1**), where it is even the second leading cause of cancer deaths. In general, the causes of the disease are poorly understood, and treatment of advanced stages is often ineffective.

The natural history or development of colorectal cancer has been well-documented to be a relatively slow process starting from easily treatable abnormalities.[2] Population screening is therefore widely believed to provide an important opportunity for disease prevention and aversion of disease-related mortality. Consequently, in the last two decades, many programs have been implemented worldwide.[3,4] However, although benefits of screening are well-established, the performance of screening programs is often suboptimal and even cause for concern.[5]

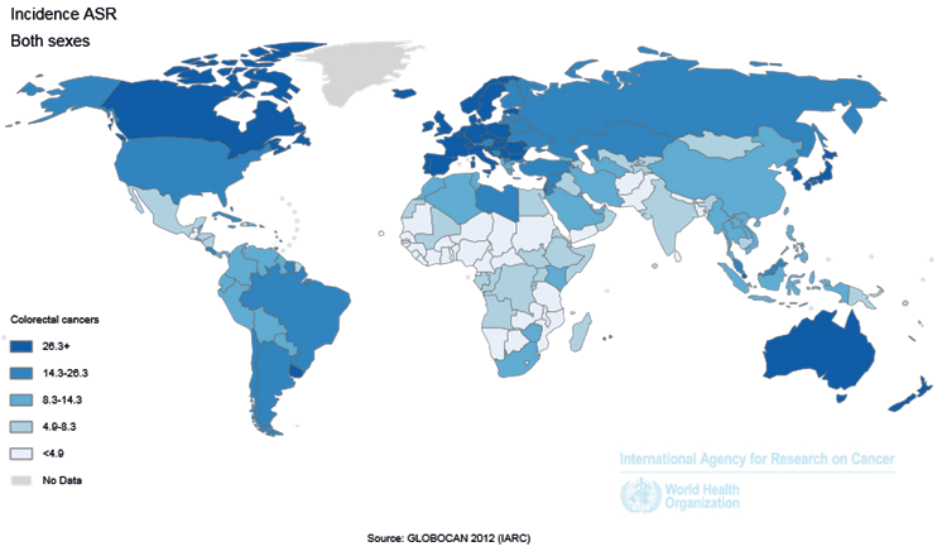


Figure 1.1 Global colorectal cancer incidence.[173]

The purpose of this thesis is to further advance the knowledge on population-level effects of screening, the best screening tests, the importance of program performance indicators for key screening outcomes, and specific questions related to personalized screening. Before addressing the central questions of this thesis in Part II-IV, in Part I, we will first provide more background information on the history of public health until present, and on colorectal cancer screening in particular. More details on methods, microsimulation modeling, are provided in Chapter 2. Finally, in Part

V, we will conclude this thesis by discussing the most important findings, strengths, short-comings, and suggestions for future practice and research.

This research is part of the public health discipline, the overarching aim of which is the "organized promotion of health and the prevention and treatment of diseases".[6]

HISTORICAL CONTEXT

In the early 20th century, medical screening has emerged as a relatively young branch on the history of public health. The origens of public health go back at least 4000 years.[6] Already in 2000 before the common era (BCE), African and Asian communities started to make use of facilities such as fresh water supplies and sewerage systems. Ancient Jewish scriptures from around 500 BCE promoted disease control measures such as isolation (quarantine) for certain infectious diseases. Also around 500 BCE, Hippocrates first inspired Greeks and Romans to adopt healthy lifestyles (diet, exercise), to strategically allocate cities near fresh water and clean soil, to build aquaducts, bathing houses, underground sewerage, and public hospitals.[7] However, despite such early milestones for public health, a true understanding of disease causes allowing for effective public health interventions, like primary prevention and screening, was lacking.

European middle ages have been characterized to a large extent by little structural advancements in public health.[6] Many of the Roman hygienic establishments were destroyed or lost to decay. Cities no longer employed active sewerage systems, sanitary workers, public health care facilities, or public health administrations. No wonder in hindsight, medieval cities were often plagued by outbreaks of infamous infectious diseases such as black death, small pox, dysentery, leprosy, and influenza.[8] Epidemics could only be controlled on a local ad hoc basis, but prevention was impossible at this stage in history.

It was during the 17th century that groundbreaking developments took place which would ultimately lead to the defeat of prevailing infectious diseases. The city of London started in 1603 to record the numbers and causes of deaths in weekly mortality bills. These allowed early epidemiologists to discover patterns in the occurrence of disease,[9] and served for instance as evidence for the benefits of small pox inoculation (18th century) and the importance of clean living conditions and water (19th century). Also in the 17th century, Anthony van Leeuwenhoek would discover microscopy and the existence of micro-organisms, a technology which would later be applied by other pioneers to discover the bacterial origins of infectious disease (19th century). Infectious diseases had lost their mystery, and could now be prevented through vaccination and improvements in public hygiene.

The developments described above substantially improved the life expectancy in Western society, and ushered in the "*epidemiological transition*" in leading causes of death from infectious- to the current prevailing chronic diseases, such as cardiovascular disease, diabetes, and cancer (**Figure 1.2**). At this stage, early 20th century, medical screening first became a favorable method for disease control. Chronic diseases were often poorly treatable at the stage of presentation. Also, there was a general lack in understanding of the complex mix of genetic, environmental and lifestyle factors causing chronic diseases, which prohibited true prevention of disease. The relatively slow progression of chronic diseases provided a good opportunity to intervene during early stages of development.

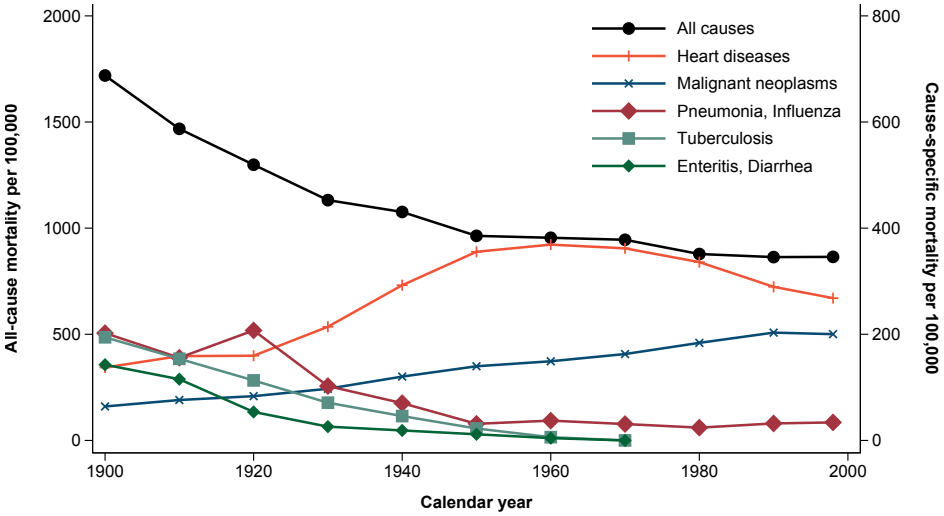


Figure 1.2 The transition in leading causes of death in the United States during the 20th century.[174]

The rise of screening, or secondary disease prevention, was spurred by the development of valid and acceptable forms of testing, and effective treatments for early stages of disease.[10] Access to health care also became more widespread during this age.

Early forms of screening included screening for syphilis at around 1950 (after already in 1906 a test had been developed to detect the presence of the causal bacteria, and treatment through penicillin became available on large-scale immediately after World War II;[11] diabetes in 1946-47 (after around 1900 insurance companies in New York already performed large-scale urine glucose testing, and treatment with insulin injections was discovered in 1923;[12-14] and cervical cancer in the 1950s (after Papanicolaou reported on the usefulness of "Pap" smear testing for detecting cervical cancers in 1928, and surgically treatable "*in situ*" lesions by 1949).[15,16] One of the first

cancer screening centers in the United States was the Memorial Hospital in New York, currently Memorial Sloan Kettering Cancer Center.

The wide application of screening for colorectal cancer is a more recent development even still. Although Lockhart-Mummery and Dukes discovered the precancerous stage of colorectal cancer already in 1927,[17] and the first major screening studies with a rigid endoscopy ("*proctoscopy*") were initiated in the United States during the 1940s, screening with proctoscopy never became popular because of an unfavorable balance in harms versus benefits.[18,19] Proctoscopy screening improved disease incidence and survival rates, but was labor-intensive, very unpleasant, and inadequate for complete colorectal examination. Operative follow-up for more proximal lesions detected on follow-up barium enema screening often required complicated surgery with a high complication rate.

Critical breakthroughs for colorectal cancer screening were the development of alternative screening methods in the 1960s (fecal testing, colonoscopy),[20,21] and the subsequent publication of evidence from several studies showing that screening could reduce colorectal cancer mortality by approximately 15%.[22] Soon professional societies followed up on this evidence by recommending population-wide screening,[23] and screening has rapidly popularized ever since.[3]

Screening is still the preferred method of prevention for many diseases and development disorders. Cancer screening is currently recommended for cervical, breast, and colorectal cancer,[24] and in the United States, also for lung cancer in heavy smokers.[25,26] Its popularity can be attributed in part to a persistent lack in understanding of chronic disease causation,[27] the ongoing development of improved screening methods, and high costs of treatment.[28] Also, the alternative method of disease prevention, primary prevention through lifestyle changes, is often difficult to achieve, even though potentially much more effective.[29] In contrast to the early days of screening, formal criteria are now used to rigorously assess the merits of screening, and to assure that the overall benefits outweigh the inevitable harms.

SCREENING THEORY

Screening definition

An often-cited definition of medical screening was published in 1951 by the Commission on Chronic Illness. It understands screening as "the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not meant to be diagnostic. Persons with positive or suspicious findings

must be referred to their physicians for diagnosis and necessary treatment."[30] This definition also applies to the case of colorectal cancer and this thesis, on the note that some forms of colorectal cancer testing do allow for immediate diagnosis and treatment.

World Health Organization criteria for screening

A comprehensive list of principles for screening evaluation was published in 1968 by Wilson and Jungner on behalf of the World Health Organization (WHO).[31] Principles included that the condition should be an important health problem, there should be acceptable treatments and suitable tests, and that the cost should be in reasonable balance with overall health care spending. Although the Wilson-Jungner criteria have become the standard criteria for screening implementation, they have some important limitations. For example, no direct screening effectiveness evidence is required by Wilson and Jungner to assess the appropriateness. Further, they used imprecise notions like 'important' (public health problem), 'suitable' (screening test) and 'adequately' (understood natural history). In actual practice, countries often use additional or more specific criteria in the spirit of Wilson and Jungner.[32,33]

In 2008, the WHO itself has updated its criteria for screening (**Box 1.1**).[34] New criteria include the requirement of a scientific basis for the effectiveness of screening, quality monitoring and assurance, and institutionalized attention for the autonomy and well-being of participants. Despite the addition of these important elements, the updated criteria leave substantial room for interpretation and require further specification in order to become practicable. Also, the cost factor of screening is no longer included, while in practice, the cost-to-benefit ratio is becoming increasingly important for screening program evaluation due to rising health care costs and aging populations.

Operational summary of criteria for screening evaluation

Our own work to inform health policy focusses primarily on three of the above mentioned criteria for screening, namely the established effectiveness of screening, the balance of benefits and harms, and the cost-effectiveness (**Box 1.2**),[35] similar to criteria proposed by Harris et al.[36] In our decision analyses, we aim to provide the evidence needed to compare available strategies for screening in terms of each of these criteria. Relevant outcome measures include: cancer deaths averted and life-years gained for effectiveness; the number of screening examinations, associated adverse effects, and over-diagnosis for harms; and cost per (quality-adjusted) life-year gained as a measure of cost-effectiveness. Effectiveness should preferably be established in randomized controlled trials. Costs are estimable from insurer data.

Although ethical considerations are important in decision making, it is not a core focus of our work to inform screening practice. Within the Erasmus MC University Medical Center, a separate department is devoted to research on ethics of screening and health care in general. Our primary aim is to inform policy makers on pivotal health outcomes and cost. Implicitly, however, there may be ethically laden assumptions in cost-effectiveness research about the importance of health benefits, harms and costs across different time periods (e.g. current versus future) and different population subgroups (e.g. young versus old, low versus high social economic status). This is explained in some more detail in the next section.

Box 1.1 Updated World Health Organization criteria for screening

(1) The screening program should respond to a recognized need.
(2) The objectives of screening should be defined at the outset.
(3) There should be a defined target population.
(4) There should be scientific evidence of screening program effectiveness.
(5) The program should integrate education, testing, clinical services and program management.
(6) There should be quality assurance, with mechanisms to minimize potential risks of screening.
(7) The program should ensure informed choice, confidentiality and respect for autonomy.
(8) The program should promote equity and screening access for the entire target population.
(9) Program evaluation should be planned from the outset.
(10) The overall benefits of screening should outweigh the harm.

Box 1.2 Erasmus criteria for screening evaluation[35]

(1) Substantial positive health benefits. Effects established preferably in randomized controlled trials
(2) Limited adverse side-effects. Anticipated balance clarified prior to participation.
(3) Reasonable ratio between cost and benefits. Ratio stable to potential short-term developments.

Cost-effectiveness analysis

Cost-effectiveness analysis is an established method to relate cost of care to estimated harms and benefits. It serves to inform policy makers on efficient allocation of limited health care budgets. Outcomes often take the form of a ratio of the level of expenditure per unit of health benefit, where the incorporated health benefits may be adjusted for potential harms from the evaluated health service. Cost-effectiveness ratios can be used to benchmark cost-pro-benefit across different health care services and sectors.

Cost-effectiveness can be assessed either incrementally from less to more effective strategies in case of mutually exclusive choices (incremental cost-effectiveness ratio, ICER), or compared to the current intervention or "null" for all evaluated strategies (average cost-effectiveness ratio, ACER).[37] The cost per (quality-adjusted) life-year gained is one of the most common cost-effectiveness metrics in the literature. This ratio divides the estimated (incremental) cost of a health service by the estimated (incremental) benefits in terms of life-years gained. Cost is usually assessed either from a third party payer perspective, primarily including direct costs of health services, or from a societal perspective, also including (to the extent possible) costs for lost productivity, travel expenses, and other indirect costs. Fixed costs are generally not considered. Life-years gained are typically assessed by first estimating the effect of a health service on disease-related mortality, as preferably established in clinical trials, and then comparing the age of cancer deaths with the average age of other cause deaths (i.e. life expectancy in the general population). Optional quality-of-life adjustments quantify the lack in quality of life for each life-year gained; downward adjustments can be used to incorporate the harmful side-effects of services, e.g. risk of disability or pain.

It is common in practice to discount future cost and benefits as included in cost-effectiveness ratios. The principle of discounting originates from finance, where there are opportunity costs and potential risks for receiving cash flows in the future rather than today. Effectively, discounting means that immediate costs and benefits are valued higher than future costs and benefits. As a consequence, benefits at younger ages are often valued higher than benefits at older ages. The convention in medical literature is to use annual discount rates of 3%,[38] however, in some countries it is customary to use different or even differential rates for cost and benefits.[39]

Some institutions have defined acceptance thresholds for cost-effectiveness. In theory, these can be used to assess whether health care expenditure per unit of benefit is in an acceptable or cost-effective range. Acceptability thresholds vary notably by country. Within the WHO-CHOICE framework, the WHO has proposed thresholds of <1x GDP per capita for very cost-effective policy, 1-3 x GDP per capita for cost-effective policy, and >3 x GDP per capita for not cost-effective, which boils down to a cost-effectiveness threshold of approximately US$100,000 for the United States and North-West Europe. For many countries, lower acceptability thresholds have been proposed, such as the United Kingdom (£20,000),[40] the United States (US$ 50,000),[38] and the Netherlands (€80,000 for curative care, €20,000 for preventive care).[41] In practice, however, effective drugs or therapies are rarely dismissed for cost-effectiveness ratios below the WHO threshold.

COLORECTAL CANCER

Disease epidemiology

Colorectal cancer ranks among the most common and deadly forms of non-communicable disease. According to the WHO,[42] approximately 724,000 (1.3%) out of all 55.9 million deaths in 2012 were attributable to the disease (**Figure 1.3**). For reference, this is more than cervical cancer and prostate cancer combined. In Northern/Southern America and Europe, the disease is more common than in other parts of the world, causing 1.9% and 2.7% of all deaths, respectively. The only cancer that is more deadly in Western countries, is lung cancer. However, unlike colorectal cancer, lung cancer is mainly caused by a known modifiable risk factor, smoking.[43] In developing countries, infectious diseases such as HIV and malaria remain the dominant causes of mortality. This explains, together with limited healthcare budgets and accessibility, why population-wide screening for colorectal cancer is no priority for most developing countries.

In 2012, 1.4 million new colorectal cancer cases were diagnosed worldwide.[44] In the past decades, the incidence has been increasing in most parts of the world along with increasing standards of living and more Western lifestyle.[1] Although some countries, such as the United States, have managed to bend these curves,[45] generally colorectal cancer is a public health problem of increasing severity.[42]

The main risk factors for colorectal cancer include age, gender and family history of disease. Incidence increases steeply with respect to age, and men have approximately 40% higher risk of getting cancer than women (**Figure 1.4**). Persons with one first-degree relative diagnosed with colorectal cancer have, on average, 2-fold higher risk compared to average-risk, while people with 2 or more diagnosed relatives may have even 4-fold increased risk.[46] Familial risk is sometimes conferred though hereditary syndromes such as familial adenomatous polyposis and Lynch syndrome.[47] Human microbiota and their DNA have also been identified as potential risk factors for colorectal cancer,[48] however, the relationship is largely still to be revealed. Finally, during the 1990s many lifestyle factors have been associated with increased colorectal cancer risk, including tobacco, alcohol, and red meat consumption (relative risks 1.4-1.5).[49] Conversely, multi-vitamin use, aspirin use, and active lifestyle may decrease risk (relative risk 0.5-0.7).

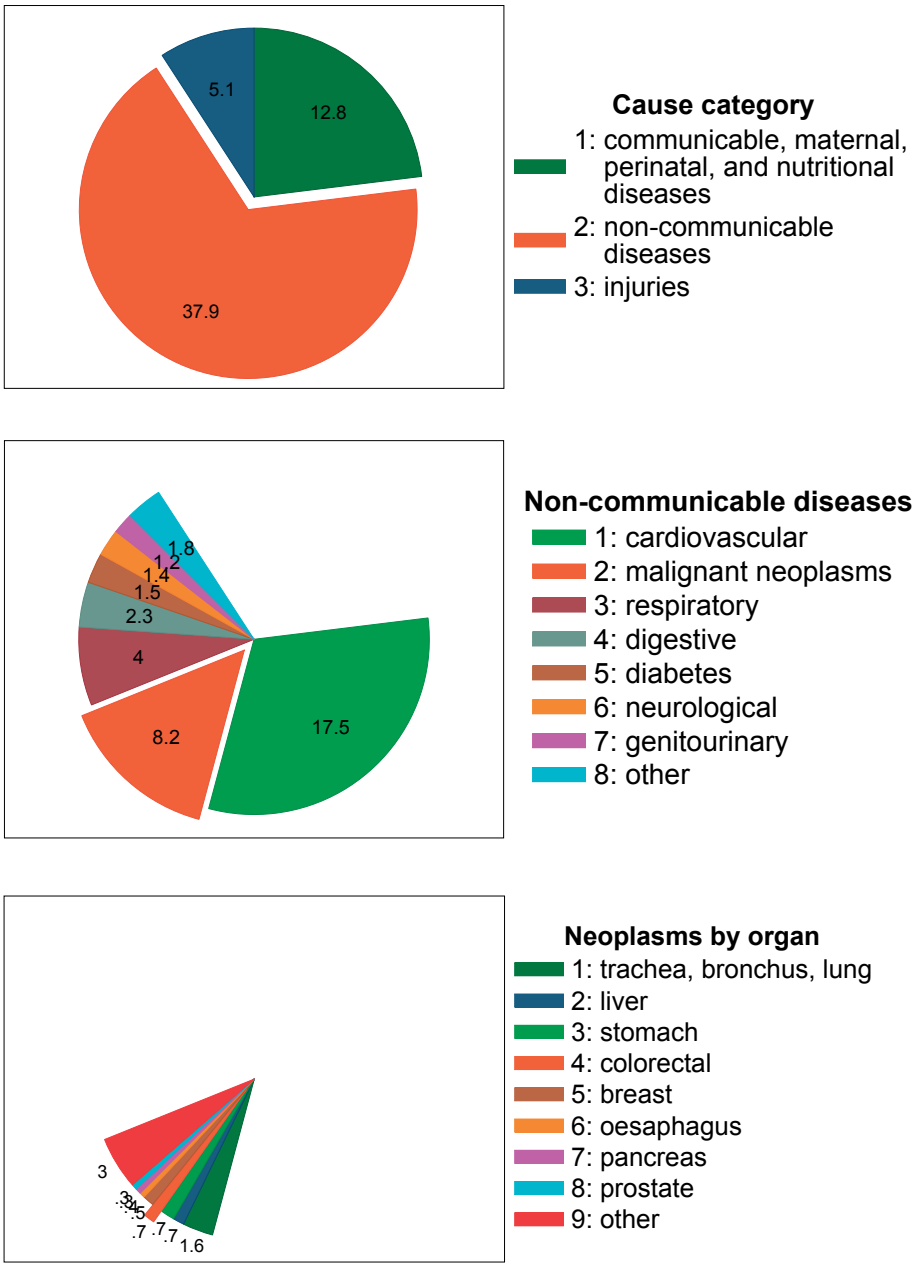


Figure 1.3a-c Causes of death, in millions.[42]

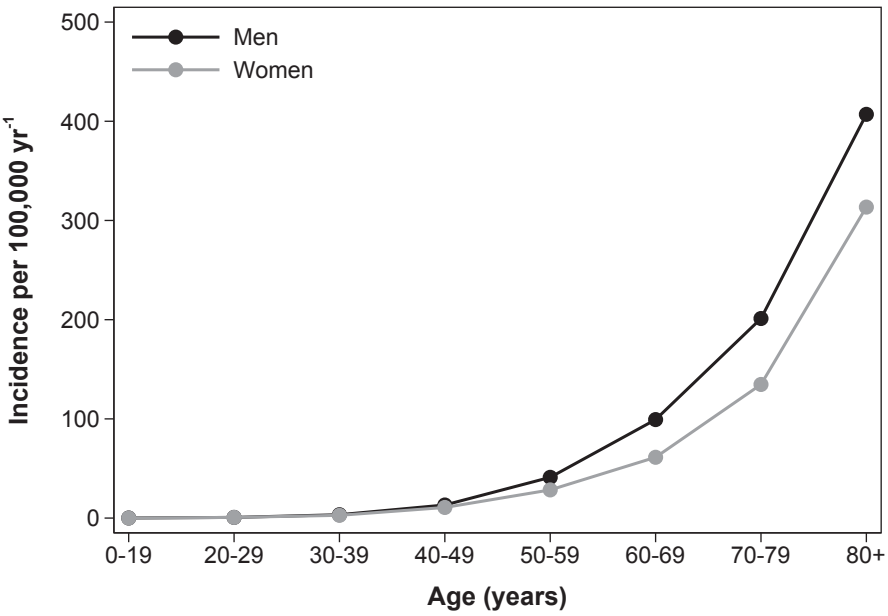


Figure 1.4 Colorectal cancer incidence by age and sex.[175]

Natural history

Colorectal cancer is detectable at an early stage of development. It has a latent preclinical phase of on average 2-4 years,[50] that can be detected by all existing screening methods (see **Screening**). Most cancers develop from benign precursor lesions or polyps visible by endoscopy and imaging. The majority of cancers (65-95%) are believed to develop through the adenoma-carcinoma sequence,[51,52] from adenomatous polyps (adenomas). A minority develops through alternative pathways, the most important one starting from sessile serrated polyps.[51]

Conventional adenomas usually have a pedunculated shape (stalked) and tubular or tubulovillous histology.[53] Some adenomas may have alternative optical features (elevated, flat, depressed) or histology (villous). Adenomas may grow in size to become more than 50mm in diameter. Generally, adenomas do not penetrate the lining of the colorectum, but when high-grade dysplasia occur they may at some point turn malignant. Approximately 30-50% of people will develop one or more adenomas throughout their life. The risk of each individual lesion progressing to cancer is much lower, with only about 4-5% of people in the United States and Europe developing the disease. It has been estimated that the average time from adenoma onset to cancer incidence, in cancer patients, is approximately 20 years,[50] leaving a substantial window for early detection and removal of adenomas before

cancer progression. High risk adenoma characteristics include size (diameter >1cm), number, histology (villous) and the presence of severe dysplasia.[54,55]

The sessile serrated pathway to colorectal cancer is believed to cause 5-33% of all colorectal cancers.[56] Serrated lesions include hyperplastic polyps without any malignant potential, sessile serrated polyps, and traditional serrated polyps. The latter two types of lesions are microscopically distinguishable from conventional adenomas by their saw-tooth configuration. Also, serrated polyps generally have different molecular features, such as more increased tumor suppressor methylation (CpG island, MLH1), BRAF gene mutations, and micro-satellite instability. The approximate average duration from onset to cancer in serrated adenomas is believed to be somewhat shorter, at approximately 15 years.[57,58]

Cancer stages differentiated in this thesis are stage I through IV according to the 5th edition of the American Joint Committee on Cancer Staging Manual (**Table 1.1**). Stage I cancers are local, invading but not penetrating the submucosal layer and muscles of the colorectum. Stage II cancers invade and potentially penetrate the outer layer of the colorectum, and may invade other adjacent organs or structures. In stage III, cancer is metastasized in one or more lymph nodes. Finally, stage IV cancers are tumors with distant metastasis. Although more recent stage classifications exist, these were not used in our studies due to unavailability of corresponding registry data (**Chapter 2**). Other types of cancer than adenocarcinoma are more rare, less well studied, and not in the scope of this thesis (carcinoids, lymphomas, sarcomas, melanomas and squamous cell carcinoma).

Table 1.1 Colorectal cancer stages according to the American Joint Committee on Cancer [a]

Stage	TNM[a]	Description
0	Tis, N0, M0	Tumor remains 'in situ', i.e. intraepithelial or invading the lamina propria
I	T1-2, N0, M0	Tumor invades the submucosa (a) or muscularis propria (b)
II	T3-4, N0, M0	Tumor invades the subserosa or pericolorectal tissues (a), or other organs and structures (b)
III	T1-4, N1, M0	Tumor affects 1-3 regional lymph nodes (a) or 4 or more (b)
IV	T1-4, N0-2, M1	Tumor affects distant organs

[a] For staging, we follow the common classification of stages as proposed by the American Joint Committee on Cancer in their 5th edition Manual for Cancer Staging,[149] which for colorectal cancer is closely related to alternative classifications such as TMN[150] and Dukes' staging.[151]

[b] In TNM classification, T reflects the invasiveness of the primary tumor, N the number of lymph nodes affected, and M the metastasis to distant organs.

Patients

There are accepted clinical guidelines on treatment of colorectal cancer patients (see **Treatment**).

Patients with adenomas detected in screening are not monitored as intensively as cancer patients, but they are examined more often than average-risk persons without detected adenomas. Surveillance guidelines in Europe and United States are largely similar, with the distinction that European guidelines indicate annual colonoscopy surveillance for some very high-risk patients (large adenoma number/size), while US guidelines generally recommend three-year intervals for such patients.[59,60] Although there is agreement among experts regarding the use of surveillance colonoscopy in patients with removed adenomas, in actual practice guidelines may often not be complied with: multiple surveys have exposed that many physicians are not familiar with or abiding by guidelines.[61-64]

In this thesis, screening participants are not considered patients unless they are diagnosed with a disease requiring close monitoring, medication, or treatment. In the Introduction and Discussion of the thesis we avoided using the term "patients" for average-risk screening participants. In the intermediate parts, we did use the term patients for screening participants according to American journal standards

Treatment

There are three dominant methods for cancer treatment. Surgery was first applied effectively for cancer treatment around 1870 by Joseph Lister. Radiotherapy was introduced soon after Wilhelm Röntgen discovered x-ray in 1895.[65] Finally, around 1950 chemotherapy was discovered as another method for treatment of leukemia by Sidney Farber and others.[66] Most colorectal cancers are still treated with one or more of these options.

The intensity, cost, and effects of treatment are related to the stage of disease. According to the National Comprehensive Cancer Network guidelines,[67] stage I cancers should generally be treated using surgery. For stage II cancer adjuvant chemotherapy and radio treatment may be recommended. More advanced cancers are treated using increasingly intensive combinations of surgery, chemotherapy, and radiotherapy. Precursor lesions, including adenomas with high-grade dysplasia, can usually be removed using "*polypectomy*", or the targeted resection of cell tissue using endoscopy with snare.

Population-wide screening programs can be effective only when treatment facilities are available and accessible for a large majority of the population. In most Western countries today, the availability of facilities for diagnosis and treatment of colorectal cancer is not a limiting factor. This is underscored by the expenditure levels for cancer treatment: in the United States alone, the treatment expenses for

colorectal cancer in patients aged 65 or older were estimated at US $14 billion for 2008,[68] and overall cancer treatment expenditures in 2012 were US $ 88 billion.[69] For certain population subgroups in Western countries, healthcare accessibility is still an important problem, mainly because of financial barriers.[70] In Europe healthcare accessibility generally is higher than in the United States. However, even in Europe, a significant proportion of the population cannot afford health care.

In developing or third-world countries the resources and facilities available for diagnosis and treatment of any kind of disease are more limited, and devoted to more urgent healthcare problems.[71,72] The Asia Pacific Working Group on Colorectal Cancer, identified healthcare access as one of the potential hurdles for implementation for screening in Asian countries in 2005: "Health-care systems and health insurance cover only a minority of people. Furthermore, access to health-care facilities is limited in many rural areas and communities of low socioeconomic status."[73] Since 2008, however, screening has been recommended for high incidence countries in the region.[74]

Treatment accessibility is closely linked to cost. Good estimates of the costs of treatment are often unavailable, and quickly become outdated when new treatments enter the market. In this thesis, we used 1998-2003 reimbursement data from the Medicare and Medicaid Services in the United States to approximate cost of screening and treatment. We adjusted payments for general inflation as measured by the Consumer Price Index. In our analyses, we also include patient time costs. The data suggest that treatment costs approach $75,000 per annum in advanced-stage cancer cases (**Table 1.2**). This may be an underestimate given approval by the United States Federal Drug Agency of several novel drugs after 2003, such as bevacizumab and cetuximab (2004), and more recently, aflibercept (2012), ramucirumab (2015) and trifluridine / tripiracil (2015).

Inverse to the increasing intensity and cost of treatment for more advanced disease stages, 5-year survival rates deteriorate with each disease stage, from 89.9% in stage I, to 70.5% in stage II-III, to 12.9% for stage IV.[75] For many of the approved expensive chemotherapeutic agents in the United States available for treatment of advanced-stage colorectal cancer patients, the additional median survival benefits compared to established treatment regimens are in the order of on average 1-2 months.[76-79] Average survival benefit is often not reported.

Table 1.2 Treatment and screening cost estimates.[152 a]

	Disease stage	Third party payer	Societal perspective	Patient time (hr)
Cancer care				
Initial care	I	29,943	34,180	243.5
	II	41,322	45,559	243.5
	III	50,383	54,620	243.5
	IV	65,791	70,028	243.5
Continuing care	I	2,383	2,723	19.56
	II	2,220	2,561	19.56
	III	3,174	3,515	19.56
	IV	9,839	10,180	19.56
Terminal care CRC	I	53,677	58,598	282.8
	II	53,525	58,446	282.8
	III	56,399	61,320	282.8
	IV	75,692	80,613	282.8
Terminal care OC	I	13,225	18,146	282.8
	II	11,567	16,488	282.8
	III	15,303	20,224	282.8
	IV	41,090	46,011	282.8
Screening				
gFOBT	-	5		0
FIT	-	26		0
COL w/o polypectomy	-	585	917	8
COL w/ polypectomy	-	762	1133	8
Complication of COL	-	5966	6245	16

Abbreviations: hr = hour; CRC = colorectal cancer; OC = other causes; gFOBT = guaiac-based fecal occult blood test; FIT = fecal Immunochemical test; COL = colonoscopy.

[a] Care was divided in four phases. The terminal phase takes precedence over the initial and continuing phase. The terminal phase reflects the last 12 months of life. The initial phase reflects the 12 months following diagnosis for persons who survival for more than 12 months (if survive for ≤12 months, person only experiences the terminal phase). The continuing phase is the time between the initial phase and the terminal phase for persons who survive for more than 24 months. All costs in **Table 1.2** have been annualized. Persons who survive only a fraction of a given phase of care should be assigned only a fraction of the costs of that phase.

Screening

Compared to screening for cancer in other organs, there is a relative wealth of screening options for colorectal cancer. Already in the 19th century, rigid endoscopes were developed to inspect rectum and part of the sigmoid colon.[80] X-ray (barium enema) imaging was used to examine the whole colon in patients with detected lesions. In the early 1960s, Overholt developed the flexible sigmoidoscope (FSIG),[81] which was used to endoscopically examine the rectum and entire sigmoid colon. In

1967, an internist from the United States recognized that colorectal cancer patients often suffer from major rectal bleeding and developed an alternative, non- invasive test for occult blood in stool.[20] In the early 1970s, full colon examination with colonoscopy was introduced, which also allowed for the immediate removal of detected lesions.[21] Over time each of the tests have been further improved or replaced by more advanced techniques.

Currently, four types of colorectal cancer tests can be distinguished: blood-based biomarker tests, stool-based tests, endoscopic examination techniques, and radio-imaging techniques (**Table 1.3)**. There is one approved blood-based test in the United States, Sept9, which qualitatively detects the methylated Septin 9 gene (approved by the Federal Drug Agency, April 2016). Common stool-based tests include guaiac-based fecal occult blood testing (FOBT) specific for the heme component in blood, fecal immunochemical testing (FIT) specific for the globin molecule in human blood, and multi-target stool DNA testing (sDNA) adding molecular assays for KRAS mutations, NDRG4 and BMP3 gene methylation, and Beta-actin hemoglobin. Endoscopic tests include FSIG and colonoscopy, and also capsule endoscopy or capsule endoscopies. Finally, radiography-based tests include CT colonography and the more dated double-contrast barium enema (DCBE).

Table 1.3 Performance of common colorectal cancer screening tests

Type	Subtypes	Evidence	Test Sensitivity [a]	Incidence Reduction	Mortality Reduction	Reference
Blood	Septin 9	Indirect	.48	-	-	[153]
Stool	gFOBT (Hemoccult II)	RCT (4)	.25-.52 [c]	.00-.20 [c]	.09-.32	[118,154-159]
	sFOBT (Hemoccult SENSA)	Indirect	.64-.80	-	-	[160]
	FIT	Observational	.73-.88	-	.10-.22	[104-106]
	mt-sDNA (Cologuard)	Indirect	.92	-	-	[161]
Endoscopy	FSIG	RCT (4)	.95	.18-.26	.21-.31	[117,119,120,162,163]
	Colonoscopy	Observational	.95	.48-.91	.68-.88	[163-169]
	Capsule endoscopy	Indirect	.88	-	-	[170]
Imaging	DCBE	Indirect	.48	-	-	[171]
	CTC	Indirect	.67-.94 [b]	-	-	[104]

Abbreviations: gFOBT = guaiac-based fecal occult blood testing; sFOBT = high-sensitivity gFOBT; FIT = fecal immunochemical test; mt-sDNA = multi-target stool DNA test; FSIG = flexible sigmoidoscopy; DCBE = Double-contrast barium enema; CTC = computed tomographic colonography.

[a] Sensitivity for cancer is reported. Most stool-based tests have low sensitivity for adenomas, imaging has low sensitivity for diminutive lesions but high sensitivity for large adenomas, and endoscopy has high sensitivity for all adenomas (>75%).[163]

[b] Sensitivity for adenomas over 10 mm in diameter

[c] The incidence reduction achieved in the Minnesota Colon Cancer Control Study of 20% was achieved with rehydrated guaiac fecal occult blood test slides, which are associated with higher sensitivity of approximately 90%.[172]

All of the above-mentioned screening methods have their particular advantages and disadvantages, that each individual patient may weigh differently. Stool-based tests are the most simple, least intrusive and cheapest of all colorectal cancer tests; blood-based tests may be more acceptable to some people, and can be completed during routine patient examinations by primary care physicians; CT imaging is a more sensitive and protective non-invasive test than the above tests; finally, endoscopy is the most sensitive for detecting both colorectal adenomas and cancer, and allows for immediate treatment of adenoma patients. All tests except primary colonoscopy require follow-up of positive results with colonoscopy. Disadvantages of the stool-based and blood-based tests are that they have to be repeated frequently; CT colonography requires inconvenient cathartic preparation and may impose risks associated with low-dose radiation; colonoscopy disadvantages include price, risk of perforation, and inconvenience.

Not surprisingly given the different test features, the actual observed acceptability of tests has been suggested to differ across settings and cultures.[82] However, the general pattern across randomized clinical studies comparing non-invasive tests with endoscopy is that non-invasive methods are preferred by most people.[83-88] Paradoxically, survey data from the U.S. suggest that, in practice, Americans tend to use colonoscopy more often than FOBT.[89] Part of this apparent paradox for the United States may stem from the differences in short- and long-term preferences. It is uncertain from existing evidence whether people would have been willing to comply with high-frequency screening methods for longer periods. Some programs have suggested reasonable compliance rates with FIT across up to 4 rounds of testing.[90-92] Although high adherence for innovative testing methods such as sDNA and Sept9 is less well-established, these tests may have the potential to be more acceptable to some people.[93]

In the United States, approximately 60% of the population reported being up-to-date with screening recommendations for any type of test.[94] Regionally, programs with mailed outreach have reported screening rates of over 80%.[95,96] In most other countries, uptake rates are lower. For example, colonoscopy-based screening programs in Germany and Poland have population uptake rates of only approximately 15%.[97] FOBT programs also do not necessarily have high adherence rates. For example, in Belgium reported uptake rates are less than 10%.[5] On the other end of the spectrum are Netherlands, Finland, and the Basque country in Spain, where uptake rates are higher than 60%.[5,98] Overall, data suggest that the acceptability of the existing screening methods for colorectal cancer may be lower than that for breast and cervical cancer, which may be attributable in part to gender differences in preferences and, for some settings, to high costs of colorectal examinations.[99]

Costs differ substantially across screening tests. In the United States, a single low-sensitivity guaiac-based FOBT (gFOBT) is reimbursed at approximately US $5, a single FIT is reimbursed at approximately US $26, colonoscopy without polypectomy is reimbursed at on average US $585, and colonoscopy with polypectomy at US $762 (**Table 1.2**). Similar to treatment, we used Medicare reimbursement and co-payment rates to approximate costs. Indirect (societal) costs of colonoscopy are relatively higher for invasive screening modalities due to the cathartic preparation, traveling, and more substantial procedure time. It is not clear to what extent fixed costs are incorporated in per-test reimbursements. Overhead costs for screening programs may be substantial, although less so for opportunistic than for organized programs (see **Programs**).[100]

Evidence for the effectiveness of screening tests is of variable quality, and often indirect. To our knowledge, Sept9 is the only used blood test for colorectal cancer. The approval by the United States Federal Drug Agency was based on evidence suggesting it may detect around 70% of cancers, with a specificity of greater than 80%.[101] More recent evidence suggests a lower sensitivity of only 50% for Sept9 (**Table 1.3**). For stool-based tests, randomized controlled trials have been conducted only for gFOBT with Hemoccult II (Beckman Coulter Inc.), suggesting that this may reduce colorectal cancer-related mortality by approximately 9-33% (intention-to-screen) (**Table 1.3**). The effects of current stool-based tests are generally deemed higher than those of gFOBT,[102,103] because these tests have better diagnostic performance characteristics than classical guaiac-based tests.[104] For FIT, recent population-based studies have suggested that the mortality reduction for screening may be as high as 62%, although the observed corresponding population-level effects of programs were still modest due to low population participation (**Table 1.3**).[105,106] For endoscopic tests, sigmoidoscopy is the only test with effectiveness estimates from trials. The estimated mortality reduction across 5 trials was 21-31% (**Table 1.3**), with higher per-protocol effects and effects for the distal end of the colorectum (45%).[107] For colonoscopy, direct estimates of the effectiveness are available only from observational studies with inherent weaknesses (primarily selection bias). These studies suggest that screening colonoscopy may reduce cancer-related mortality by even higher percentages >50% (**Table 1.3**). The effectiveness of capsule endoscopy has not been assessed directly, however, the estimated diagnostic performance is for large adenomas and cancer of 88% is close to that of colonoscopy. Likewise, no direct evidence for the effectiveness of radio-imaging exists, but sensitivity for large adenomas is comparable to colonoscopy (**Table 1.3**).

Given the high present level of treatment costs, multiple independent modeling studies have indicated that colorectal cancer screening is very cost-effective.[108,109] For the Dutch screening program, Wilschut and colleagues evaluated a number of

different screening strategies varying in terms of test, interval, and age range. All of the evaluated strategies had ACERs of less than €20,000 per life-year gained.[110] Studies for the United States Centers for Medicare and Medicaid Services from 2010 also indicated that colorectal screening, independent of screening modality, is highly cost-effective, with ACERs ranging from less than US $0 up to $14,000.[111]

Screening for colorectal is recommended in most Western countries. Expert panels across and within countries differ in the types of tests included in the recommendations for screening. The Council of the European Union endorsed only annual or biennial screening with sensitive guaiac-based fecal occult blood testing (sFOBT) between ages 50-74 years,[112] which was reconfirmed more recently by a pan-European expert panel.[113] In the United States, the American Cancer Society, the United States Multi-Society Task Force on Colorectal Cancer, and American College of Radiology recommended screening from age 50 with either colonoscopy at 10 year intervals, FSIG, DCBE and CTC at 5 year intervals, sFOBT or FIT with 1 year intervals, and sDNA with unspecified interval;[103,114] the United States Preventive Services Task Force recommended no screening after age 75, and recommended the same tests except for DCBE, which they replaced with Sept9; the American College of Gastroenterologists favors colonoscopy for screening.[115] For South-East Asia and Australia, the Asian Pacific Working Group on Colorectal Cancer restrict recommended strategies for screening to FIT.[74]

RECENT DEVELOPMENTS

Programs

Many screening programs have been initiated over the last two decades.[4] Although available information on screening programs is incomplete, fully implemented organized programs are known to exist at least in Croatia, France, Slovenia, parts of the UK (England, Wales, and N. Ireland), several Canadian provinces, Israel and Japan. Organized programs are being rolled out in many other countries, including Australia, Belgium (since 2013), Denmark (since 2014), Finland (since 2009), Poland (since 2000), and the Netherlands (since 2014). Many other countries, including Russia, are running pilot programs. Some countries have opted for opportunistic screening programs, where screening may be promoted but is not offered by the government. Countries with opportunistic programs include Austria, Germany, Czech Republic and Greece in Europe, and Uruguay and the United States in the Americas. In Europe, no screening activity is currently known to exist in Bulgaria, Estonia, Hungary, and Romania.

The Dutch screening program as initiated in 2014 will roll out screening over a period of five years. By 2019, biennial FIT screening will be offered to all men and women aged 55 through 75 years. Screening is performed with a new kind of FIT, the FOB Gold (Sentinel Technologies Inc.), which will be sent and collected via mail. Periodical monitoring and modeling is used to control positivity and colonoscopy referral rates for the test.[116] The estimate is that the program will eventually prevent 2400 deaths from colorectal cancer per year. For the Netherlands, initial program performance looks promising, with first year adherence exceeding 70%.[98]

Current topics in research

Recent research related to colorectal cancer screening has focused on several areas for improvement of screening programs.

First, as already addressed to some extent, studies have looked at the question of the effectiveness of stool-based versus endoscopic testing for colorectal cancer. In the past five years, four studies have been published which evaluated the benefits of FSIG screening.[117-120] In 2015, the first data regarding population-level effects of FIT were published.[105,106] More definite answers on the question of effectiveness in colonoscopy versus stool-based testing will come from randomized clinical trials. Trials have been initiated to evaluate 10-15 year benefits of colonoscopy and FIT screening, but results are not expected before 2025.[86,121,122] In advance of trial results, in this thesis, we used modeling to compare the effectiveness of fecal colorectal cancer testing with colonoscopy screening (see **Research questions**).

A second major development, non-exclusive to colorectal cancer screening, is the utilization and evaluation of performance indicators. In the United States, physician reimbursement will be made dependent on quality indicator scores, some of which still are to be validated. An obvious and critically important determinant of outcomes for screening programs is screening adherence. Already in 2000, cost-effectiveness studies have suggested that the outcomes of FOBT screening may be highly sensitive to adherence rates.[123] As we mentioned in the previous section (see **Screening**), multiple trials from the past 10 years have found that initial test adherence may be higher for stool-based testing than for colonoscopy.[83-88] Only very recently, the first population-based studies and trials have reported overall FIT or sFOBT compliance rates over up to 4 rounds of testing, with cumulative adherence rates of around 50%.[90-92] Translation to primary cancer-related outcomes has not been performed. In this thesis, we present novel randomized clinical trial data from the U.S. with adherence and outcome data for up to 7 rounds of sFOBT screening versus colonoscopy. Modeling was used to assess and compare the long-term corresponding benefits.

Colonoscopy quality is another important outcome determinant that has received a lot of attention in recent years. In the past six years several high-impact studies

have suggested that there is a strong association between the main colonoscopy quality indicator, adenoma detection rate (ADR), and patient outcomes.[124,125] Other studies have looked at alternative quality indicators such as adequacy of bowel preparation,[126,127] colonoscopy withdrawal time,[128,129] and colonoscopy completion (or cecal intubation) rates.[127,130] The American College of Gastroenterology Task Forces on Quality in Endoscopy has reviewed currently available data and published an update on what quality indicators to use for colonoscopy to assure satisfactory screening outcomes.[131,132] Primary recommended quality indicators include, both the ADR (target value <25%), cecal intubation rate (>95%), and adequate follow-up recommendation rate (>95%). This thesis includes a report on the estimated relationship between observed ADR adenoma variation and colorectal cancer screening benefits, risks, and cost.

A third development, again broader for general health care, is a shifted focus toward more personalized health services. It is recognized increasingly that every person is different, and may not benefit equally from a certain form of therapy or screening. Current colorectal cancer screening guidelines only take into consideration persons' age, their history of adenomas, and family history of disease and polyposis syndromes, but no other known risk factors for colorectal cancer. In the past 5 years, several studies have been published focusing on prediction of risk of neoplasia from personal characteristics including age, gender, BMI or waist circumference, family history, and smoking.[133-135] Hemoglobin levels from quantitative stool tests have also been proposed for risk estimation.[136,137] Most risk prediction tools have not been externally validated. Implications of risk scores for optimal screening strategies are unknown, the assessment of which may require modeling.[138] In this thesis, we investigated the relevance of general health status and screening status for the question whether to screen elderly patients, as well as more intensive examination of adenoma patients (discussed below).

Finally, there is an increasing amount of attention for management of patients with colorectal adenomas. It is well-known that adenoma patients are at higher risk for colorectal cancer,[139] however, scant data are available to compare cancer outcomes for different surveillance strategies. Current recommendations lean heavily on evidence from the 1990s and early 2000s with few cancer outcome points.[54,140] There are concerns that with improved colonoscopy quality, the guidelines may no longer be appropriate.[141] In recent years, new pooled data have been published from older studies on adenoma recurrence risks with different surveillance intervals and types of baseline adenomas.[54,55,142] A Dutch study has compared adenoma findings for two intervals in patients with a family history,[143] and an American study looked specifically at surveillance benefits for elderly patients.[144] Finally, a Norwegian study has used population-based registries to compare colorectal cancer mortality risk

of patients with low- or high-risk adenoma to the general population (in a non-screening setting), finding differences of maximum 40% between high- and low-risk patients.[145] Despite these new data, evidence is still too fragmented to adequately inform clinical guidelines. In Europe, several countries are accruing patients for a large-scale, long-term randomized clinical trial, named the European Polyp Surveillance (EPoS) study.[146] In this thesis, we used modeling to estimate the effectiveness and cost-effectiveness of surveillance strategies to be evaluated in EPoS.

RESEARCH QUESTIONS

The remainder of this thesis has three main parts (Part II-IV), which focus on the estimation of population-level impact of screening (Chapter 3-5), the importance of various effect determinants in actual screening practice (Chapter 6-9), and the potential for more personalized test strategies in specific high-risk patients (Chapter 10-11). Personalized screening is interpreted here broadly as taking into account for screening regimens, any other factors than age. Adenoma surveillance is presented under the umbrella of personalized screening on the basis that adenoma findings are taken into account for recommended strategies of examination.

For each of the above chapters, the corresponding research question is formulated below. All questions were addressed using microsimulation modeling. The microsimulation model will be presented in detail in Chapter 2. Some modeling studies were partly informed by novel empirical data. Where applicable, these data are presented in the chapters listed below.

(1) What is the expected impact of achieving 80% screening rates by 2018 on CRC incidence and mortality in the United States? (Chapter 3)
(2) How does currently available colonoscopy capacity in the United States compare to estimated need under a national screening program? (Chapter 4)
(3) Which fraction of CRC mortality in the United States is attributable to nonuse of screening? (Chapter 5)
(4) What are the estimated benefits of colonoscopy versus sFOBT screening with observed adherence rates from the National Colonoscopy Study? (Chapter 6)
(5) How does observed ADR variation influence screening benefits, harms and costs? (Chapter 7)
(6) How do FIT and colonoscopy screening benefits compare at different levels of ADR? (Chapter 8)
(7) What is the outcome effect of increasing time to diagnostic colonoscopy following a positive fecal colorectal cancer test result? (Chapter 9)

(8) Up to what age should screening for colorectal cancer be considered in elderly unscreened patients? (Chapter 10)
(9) How effective and cost-effective is currently recommended surveillance of adenoma patients compared to less intensive surveillance or screening? (Chapter 11)

SUPPORT

The present work was conducted as part of the research consortium Population-based research optimizing screening through personalized regimens (PROSPR) funded by the National Cancer Institute (NCI) within the United States National Institutes of Health.[147] PROSPR is a program with the scientific goal of supporting research to better understand how to improve the screening process (recruitment, screening, diagnosis, referral for treatment) for breast, colorectal, and cervical cancer. NCI has funded seven research centers and one statistical coordinating center. The specific aims of our collaboration with Kaiser Permanente Northern and Southern California were to compare the effectiveness of a FIT program and colonoscopy screening for colorectal cancer, to study the balance of benefits and harms, and to conduct exploratory studies to inform future research to optimize screening programs.

Another important source of financial support was the Cancer Intervention and Surveillance Modeling Network (CISNET) funded by NCI.[148] CISNET is a consortium of NCI-sponsored investigators who use statistical modeling to improve the understanding of cancer control interventions in prevention, screening, and treatment and their effects on population trends in incidence and mortality. These models can be used to guide public health research and priorities, and they can aid in the development of optimal cancer control strategies.

This thesis was supervised by Dr. Marjolein van Ballegooijen, Dr. Iris Lansdorp-Vogelaar, and Prof. Harry de Koning from the Erasmus MC Department of Public Health. The sub-department involved in screening evaluation has an esteemed reputation in informing both national and international screening recommendations for various cancer types, including breast, lung, colorectal, prostate cancer, esophageal, and cervical cancer.

Chapter 2

Microsimulation Screening Analysis

The ruling paradigm of evidenced-based health care dictates that decision makers should seek for a solid empirical evidence basis to support their decision making. In actual practice, for many decisions there are no data to directly inform decisions. As we showed in the last chapter, the most rigorous study design to assess the effect of health care interventions, experimental study, is applied only selectively in actual practice. There are far more screening options and possible screening strategies then ever could be evaluated in trials. Moreover, in colorectal cancer screening, the most important disease outcomes (e.g. death from disease) occur only rarely, which means that practitioners are often forced to look at intermediate outcomes that are more difficult to interpret (e.g. adenomas). Finally, empirical studies may span many years, which means that there is inherently a lag in informed-decision making if decisions are based entirely on empirical basis. These kinds of limitations combined with advances in computer science have stimulated researchers since the 1980s to look for inventive new ways to study diseases and health care interventions.

Microsimulation modeling is an established method to inform policy decisions. Microsimulation models integrate existing knowledge on population demographics, the natural history of disease, risk factors, potential screening test characteristics and effects, and costs of care. The information is used to simulate screening in virtual populations similar to existing populations in terms of life expectancy and disease risk. Quality of models can be assessed by periodic validation to newly published outcome data. Validated models are invaluable tools for finding optimal strategies and policies for screening for diverse settings.

Applications of microsimulation modeling include extrapolation of empirical study findings for long-term effects, optimization of screening strategies in terms of intervals and starting and stopping ages, and evaluation of individualized strategies. The MISCAN model has been used many times to inform decision makers in screening evaluation and planning. In this thesis, we present a number of applications for microsimulation modeling. Studies were all conducted using the MISCAN-colon model. A more detailed description of the model is provided in subsequent paragraphs.

MODEL STRUCTURE

MISCAN-colon is a stochastic, semi-Markov, microsimulation model for colorectal cancer (CRC) programmed in Delphi (Borland Software Inc.). 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, surveillance after polypectomy, and treatment.

The term 'microsimulation' implies that the individuals are moved through the model one at a time, rather than as proportions of a cohort. The term 'semi-Markov' implies that MISCAN-colon, unlike traditional Markov models, does not assume annual state transitions; instead it generates durations in states, allowing for various time distributions and dependency of consecutive state transitions, and thereby increases model flexibility and computational performance. The term 'stochastic' implies that the model determines the states and corresponding durations by drawing from probability distributions, rather than using fixed values. Hence, the results of the model are subject to random variation.

MISCAN-colon consists of three modules: a demography module, natural history module, and screening module.

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.[176] The maximum age an individual can achieve is assumed to be 100 years.

NATURAL HISTORY MODULE

Transitions

As each simulated person ages, one or more adenomas may develop (**Figure 2.1**). 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 without symptoms, or be diagnosed during each stage CRC because of symptoms. After clinical diagnosis, CRC survival is simulated using age-, stage-, and localization-specific survival estimates for clinically diagnosed CRC as obtained by Rutter and colleagues.[177] 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 due either to CRC or another cause ('Demography module').

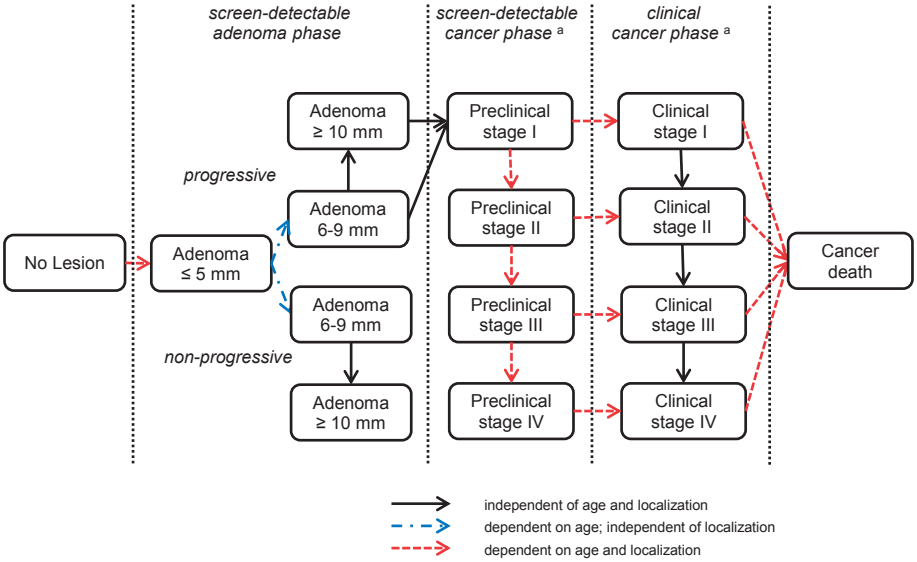


Figure 2.1 The stages of disease in the semi-Markov model.
[a] Cancer stages were based on the 5th edition Cancer Staging Manual from the American Joint Committee on Cancer[149]

Transition rates and durations

An individual's risk of developing adenomas depends on the individual's age and a personal Gamma-distributed risk index (non-homogeneous Poisson process). 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 equals the distribution of cancers as observed in SEER before the introduction of screening.[75] 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 (**Figure 2.2**).[178-187] 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 SEER data on the age-, stage-, and localization-specific incidence of CRC as observed before the introduction of screening (**Figure 2.3**).[75]

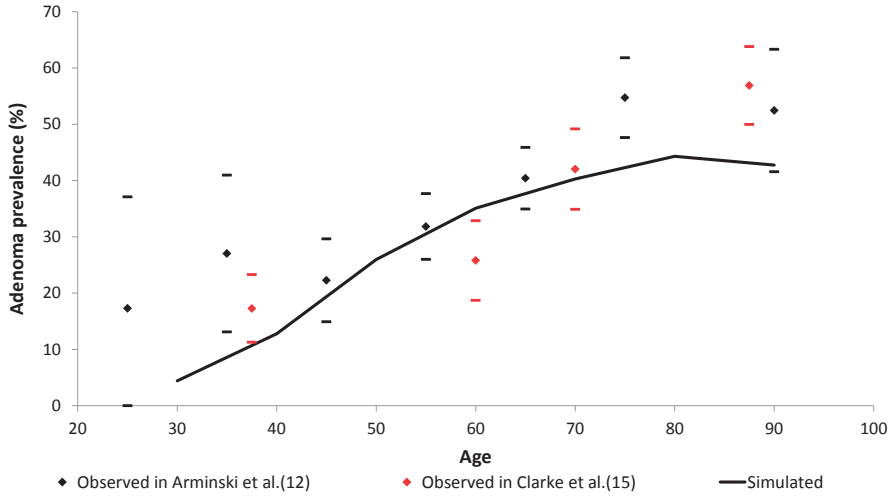
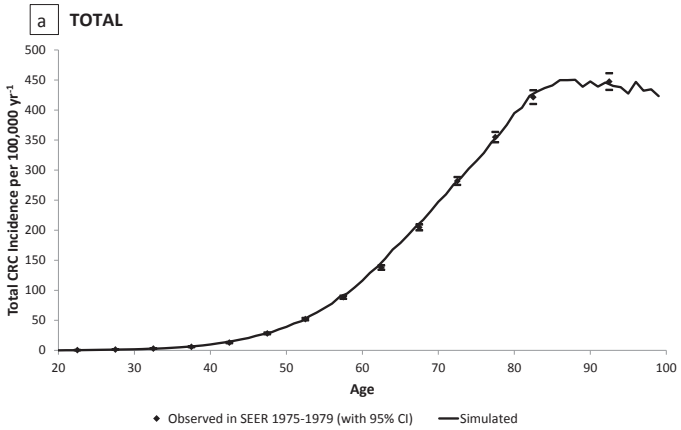
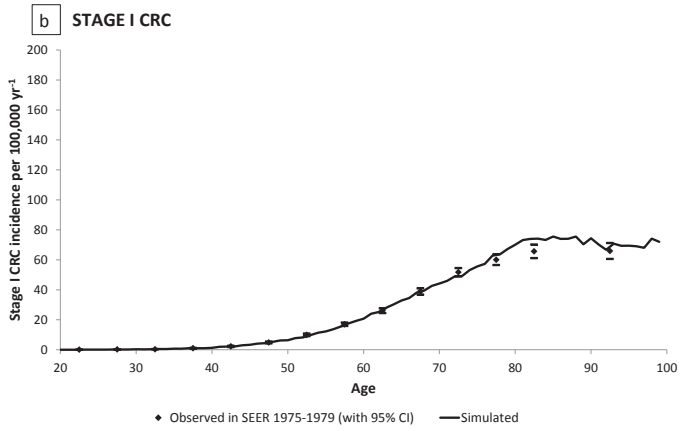
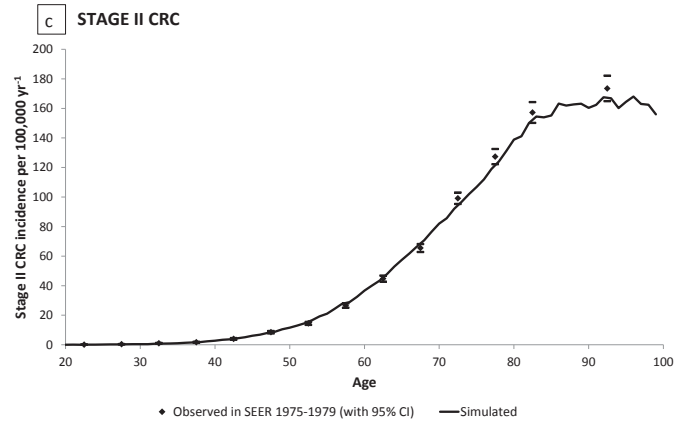


Figure 2.2 Simulated versus observed adenoma prevalence in selected autopsy studies (with 95% confidence intervals).[a]

[a] 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.


a
TOTAL
Total CRC Incidence per 100,000 yr-1
Age
Observed in SEER 1975-1979 (with 95% CI)
Simulated


b
STAGE I CRC
Stage I CRC incidence per 100,000 yr-1
Age
Observed in SEER 1975-1979 (with 95% CI)
Simulated


c
STAGE II CRC
Stage II CRC incidence per 100,000 yr-1
Age
Observed in SEER 1975-1979 (with 95% CI)
Simulated

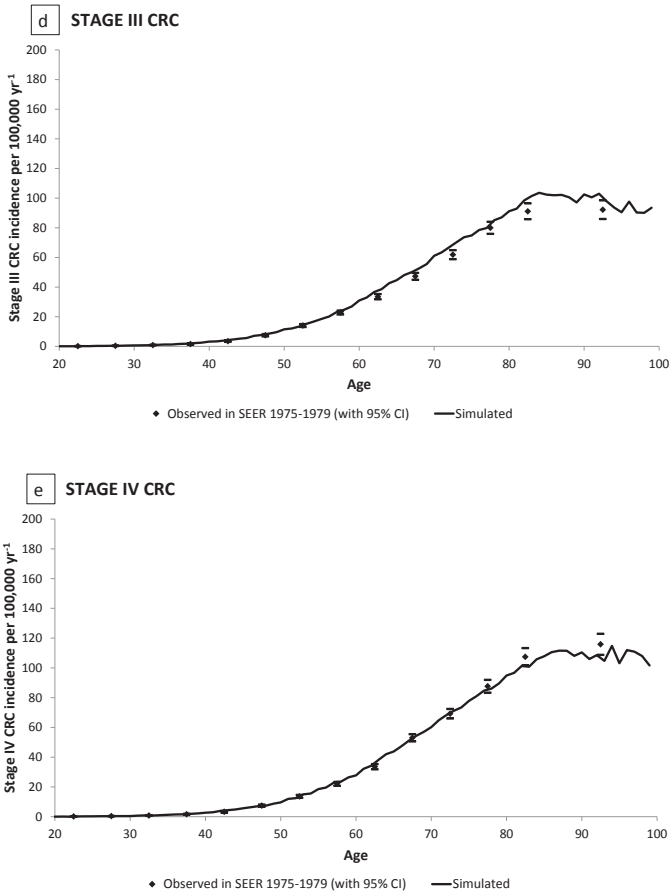


Figure 2.3a-e Simulated versus observed colorectal cancer incidence in 1975-1979 Surveillance Epidemiology and End Results program data.

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.[157,188,189] This exercise has been described extensively elsewhere.[190] The average duration from the emergence of an adenoma until progression into preclinical cancer (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 (**Figure 2.4**).[190] 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 of the disease stages within the adenoma and preclinical cancer phase, respectively, 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 the medium size, non-progressive adenoma state were calibrated to size-specific adenoma detection rates observed in a Dutch randomized controlled trial on colonoscopy screening (*not shown*). All calibrations were performed using the Nelder-Mead search algorithm to minimize deviances from observed values based on log-likelihood functions (Poisson likelihood for incidence, Binomial likelihood for adenoma prevalence, and Multinomial likelihood for cancer stages).

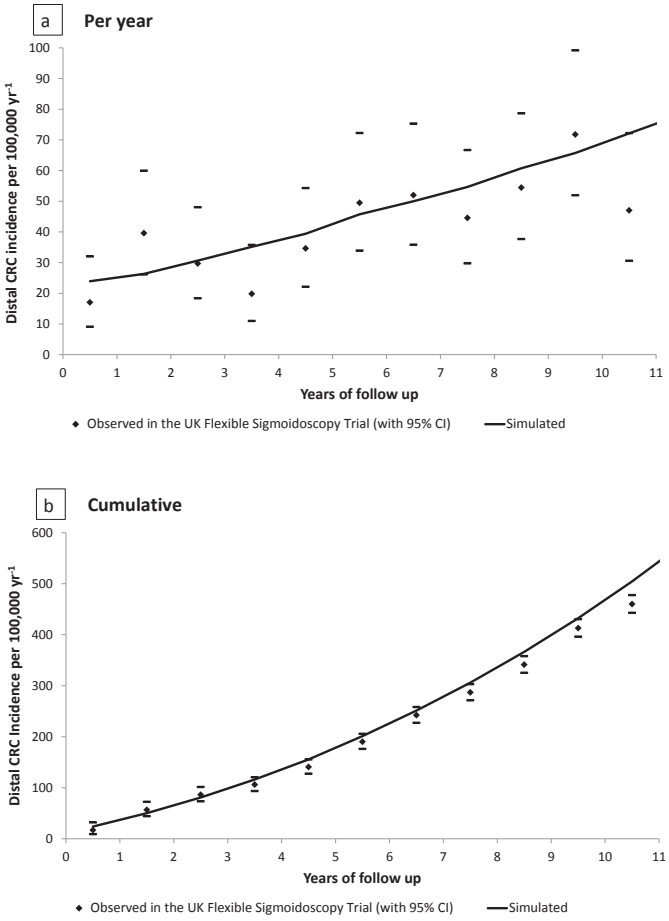


Figure 2.4a-b Simulated versus observed distal colorectal cancer incidence in the intervention group of the UK Flexible Sigmoidoscopy Trial.

SCREENING MODULE

Screening will alter some of the simulated life histories: some cancers will be prevented by the detection and removal of adenomas, while other cancers will be detected in an earlier stage with a more favorable survival. The effectiveness of screening depends on a test's assumed ability to detect adenomas and CRC. 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,[190] 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 positive health effects of screening, the model also allows for the evaluation of 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. Then, in Patient A in **Figure 2.5**, the natural history module generates an adenoma. This adenoma progresses into preclinical cancer, which, in the absence of screening, 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 black arrow. During this examination the adenoma is detected, and as a result both CRC and CRC death are prevented. Hence, integrating all 3 modules for Patient A, screening prolongs life by the amount indicated by the blue 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, indicated by the red 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, integrating all 3 modules in this patient, screening does

not prolong life but results in additional LYs with CRC care (over-treatment), as indicated by the red arrow.

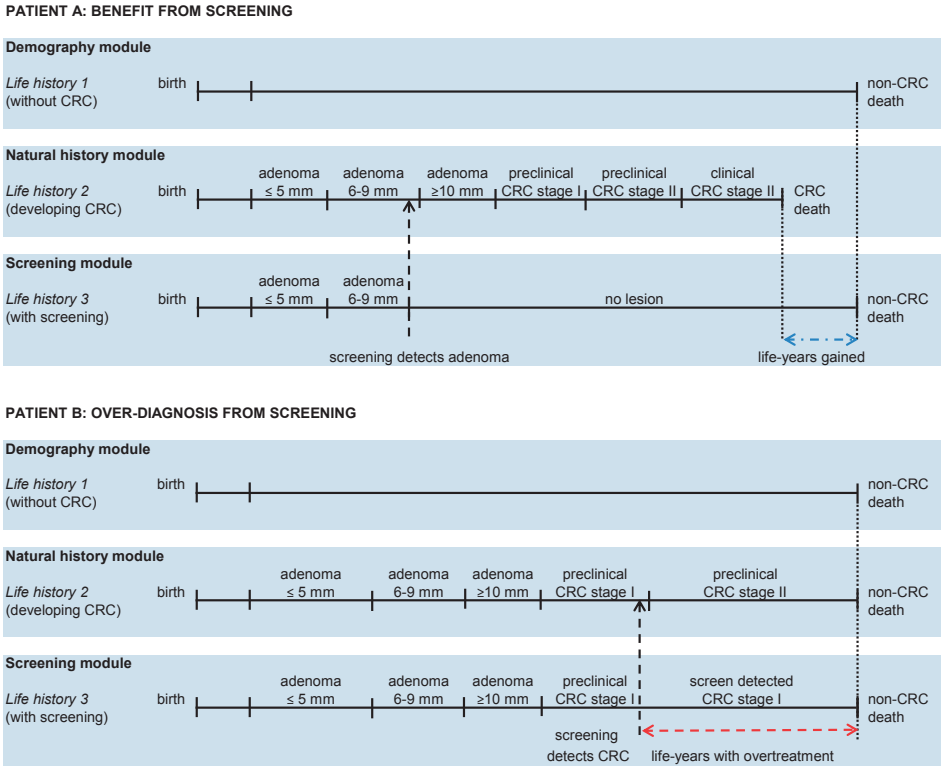


Figure 2.5 Integrating MISCAN modules for two example patients.

VALIDATION

MISCAN-colon has been validated to several randomized clinical trials. In 2004, it was validated to National Polyp Study data evaluating colonoscopy surveillance in adenoma patients.[191] In 2009, MISCAN was validated to combined data from the Minnesota Colon Cancer Control Study, the Nottingham study and Funen study of biennial gFOBT screening. Based on this study the assumptions for test performance were adjusted.[190] In 2011, we validated the model to U.K. Flexible Sigmoidoscopy Study data, and updated the model's adenoma dwell time assumptions.[50,192] Current estimated mortality risks for patients in colonoscopy surveillance are in line with long-term National Polyp Study observations (**Figure 2.6**).[193] Another validation project which uses data from the Norwegian Colorectal Cancer Prevention study on

flexible sigmoidoscopy is in progress.[194] Finally, within CISNET, efforts are ongoing to develop methodology to allow for meaningful comparisons of model differences.[195]

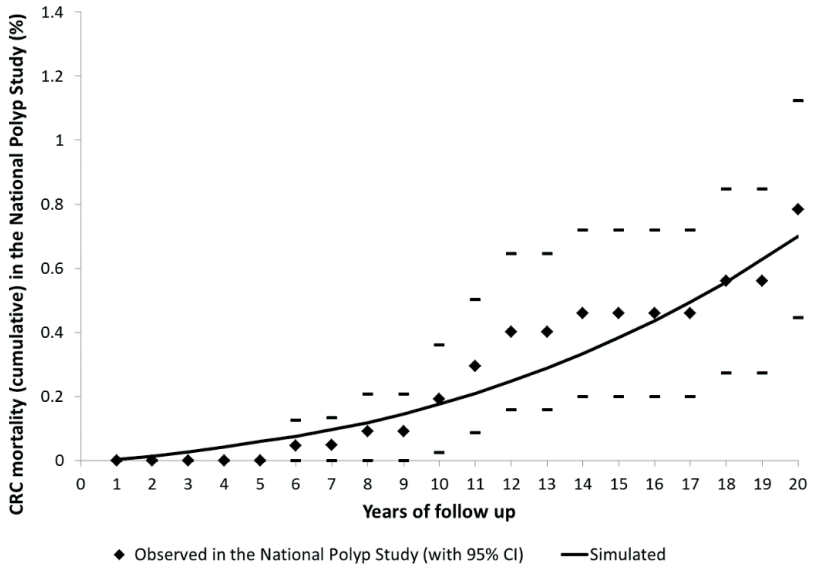


Figure 2.6 Simulated versus observed colorectal cancer mortality in the National Polyp Study

In this thesis, we validated the model to trial data from the National Colonoscopy Study data (**Chapter 6**), to observational screening data from Kaiser Permanente Northern California (**Chapter 7**), and to published adenoma surveillance data from the 1990's and early 2000's (**Appendix 11**).

Part II

Public Health Impact of Screening

Chapter 3

Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018

Article published (AltMetric = 123):

RG Meester, CA Doubeni, AG Zauber et al.
Cancer. 2015 Jul 1;121(13):2281-5.

ABSTRACT

BACKGROUND: The National Colorectal Cancer Roundtable, a national coalition of public, private and voluntary organizations, has recently announced an initiative to increase colorectal cancer (CRC) screening rates to 80% by 2018 in the United States. We evaluated the potential public health benefits of achieving this goal.

METHODS: We simulated the 1980-2030 United States population of 50-100 year-old persons using microsimulation modeling. Test-specific historical screening rates were based on 1987-2013 National Health Interview Survey data. The effects of increasing screening rates from approximately 58% in 2013 to 80% in 2018 were compared to a scenario in which the screening rate remained approximately constant. The outcomes were cancer incidence and mortality rates and number of CRC cases and deaths counts during short-term follow-up (2013-2020) and extended follow-up (2013-2030).

RESULTS: Increasing CRC screening rates to 80% by 2018 would reduce CRC incidence rates by 17% and mortality rates by 19% during short-term follow-up and by 22% and 33%, respectively, during extended follow-up. These reductions would amount to a total of 277,000 averted new cancers and 203,000 averted CRC deaths from 2013 through 2030.

CONCLUSION: Achieving the goal of increasing the colorectal cancer screening uptake in the United States to 80% by 2018 may have a considerable public health impact by averting approximately 280,000 new cancer cases and 200,000 cancer deaths within less than two decades.

INTRODUCTION

Colorectal cancer (CRC) is the fourth most commonly diagnosed cancer for both sexes combined and the second leading cause of cancer death in the United States, with an estimated 136,800 new cancer cases and 50,300 deaths in 2014.[196] Registry data from the past decade indicate that both disease incidence and mortality decreased approximately 3% per year,[197] largely due to increased use of screening.[198,199] Despite the effectiveness of screening and the availability of various screening options, only 58% of United States adults ages 50-75 years had received guideline-recommended testing in 2013.[200] Previous studies show that a substantial proportion of CRC deaths are attributable to nonuse of screening.[201,202] This rallied a recent initiative from the National Colorectal Cancer Roundtable (NCCRT), a national coalition of public, private and voluntary organizations, to aim for screening rates of 80% by 2018 in the United States.[203] However, an estimate of the potential benefits of increasing uptake by an additional 22% in terms of the number of CRC cases and deaths averted is needed to inform public discourse and policy on this initiative and to project the short- and long-term public health 'return on investment'. In this study, we used advanced modeling approaches to estimate the potential benefits in terms of new CRC cases and deaths averted from achieving the NCCRT goal.

METHODS

This study was based on men and women ages 50-100 years, simulated to match the 1980-2030 United States population in terms of their life expectancy, risk of CRC, and past and future use of screening. The analyses utilized the MISCAN-colon model (**Chapter 2**), which has been used to inform United States Preventive Services Task Force screening recommendations.[204]

Source data

Demography estimates were obtained from the United States Census Bureau;[205] overall life-expectancy was based on generational United States life tables from the Berkeley mortality database.[206] Historical use of colonoscopy, fecal occult blood tests and sigmoidoscopy in the United States were derived from 1987–2013 National Health Interview Survey (NHIS) data.[200] In 2013, 58% of the population ages 50-75 years reported up-to-date on screening. The percentages reporting up-to-date on each specific test were 54% for colonoscopy, 8% for fecal occult blood tests, and 4% for sigmoidoscopy.

Screening scenarios

In the analysis, we evaluated a scenario in which the screening rate increased linearly from 58% in 2013 to 80% in 2018, with no further increase through 2030. We compared this scenario to one in which screening rates remained constant at approximately 60%. We evaluated the magnitude of the reduction in CRC incidence and mortality rates per year during short-term (2013-2020) and extended follow-up (2013-2030). Screening consisted of a mix of colonoscopy, sigmoidoscopy and fecal occult blood testing in accordance with estimates from NHIS. Patients with a positive fecal occult blood test or sigmoidoscopy (for adenomas or cancer) were referred to diagnostic colonoscopy, and patients with adenomas detected were referred for colonoscopy surveillance according to United States guidelines.[139] Patient adherence to diagnostic colonoscopy and surveillance colonoscopy was assumed to be 80%.[207,208]

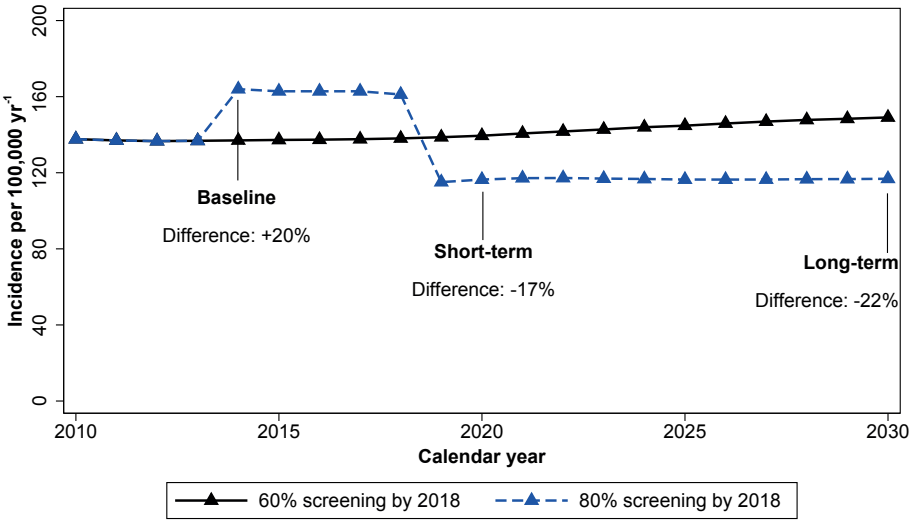
Role of funding for this study

This study was conducted within the NCI-funded Population-Based Research Optimizing Screening through Personalized Regimens (PROSPR) consortium (NCI grant U54 CA163262), which aims to conduct multi-site, coordinated, trans-disciplinary research to evaluate and improve cancer screening. The modeling for this study was also supported by CISNET (NCI grant U01 CA152959). Dr. Jemal received financial support from the Intramural Research Department of the American Cancer Society.

RESULTS

Incidence rates and avoidable new cancer cases

Under the assumption of approximately constant CRC screening levels in the United States between 2013 and 2030, the crude CRC incidence rate per 100,000 per year would increase from 137 in the first year of follow-up (2014) to 149 in 2030 (**Figure 3.1a**), due to aging of the population. If screening uptake increased from 58% in 2013 to 80% in 2018, the incidence rates (per 100,000) would decrease from 164 in 2014 to 117 in 2030. Compared to a scenario of constant CRC screening levels, '80% by 2018' would initially increase CRC incidence rates by 20% in 2014 because of early detection of CRC in previously unscreened individuals, but subsequently decrease the incidence rates by 17% by 2020, and by 22% by 2030. With an increase of the estimated population ages 50-100 years from 108 million in 2014 to 133 million by 2030, the above effects on incidence and mortality rates would result in 43,000 averted cases per year by 2030, and a total of 277,000 cases averted from 2013 through 2030 (**Table 3.1**).



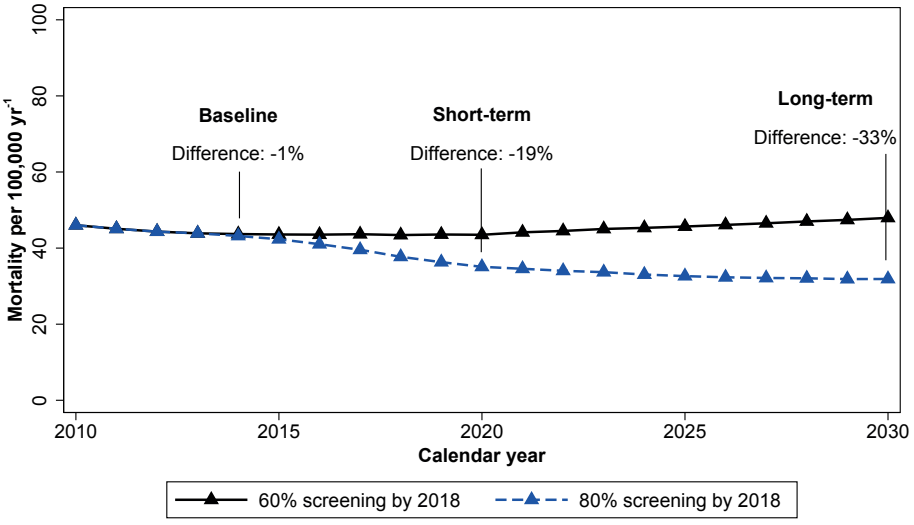


Figure 3.1a,b Crude colorectal cancer incidence (1a) and mortality rates (1b) in the United States population of age 50 and older, under two scenarios of screening uptake. In the first scenario reported screening rates remained at a constant level of approximately 60% from 2013 through 2030; in the other scenario screening rates increased from 60% to 80% by 2018 and remained constant after that.

Table 3.1 Difference in the number of CRC cases and deaths per year when achieving 80% CRC screening rates in the United States by 2018, compared to constant 60% CRC screening rates

	Calendar year				
	2014	2018	2022	2026	2030
Difference in number of CRC cases (x1,000)					
Per year	29	27	-30	-38	-43
Cumulative	29	141	28	-112	-277
Difference in number of CRC deaths (x1,000)					
Per year	-1	-7	-13	-18	-21
Cumulative	-1	-17	-60	-123	-203

Abbreviations: CRC = colorectal cancer.

Mortality rates and avoidable cancer deaths

There would be an immediate mortality benefit for 80% CRC screening by 2018. While under constant 60% screening levels the crude CRC mortality rate per 100,000 would increase from 44 in 2014 to 48 in 2030, the mortality rate would decrease from 43 to 32 with 80% screening by 2018. Thus, the relative effect of '80% by 2018' would be a 1% decrease in the CRC mortality rate by 2014, 19% in 2020, and 33% in 2030 (**Figure 3.1b**). This would translate to 21,000 averted cancer deaths per year by 2030, and a total of 203,000 averted deaths from 2013 through 2030 (**Table 3.1**).

DISCUSSION

We used microsimulation modeling to estimate the potential United States public health impact of achieving the NCCRT goal to increase CRC screening rates from just under 60% in 2013 to 80% by 2018. Our results suggest that achieving this goal may produce a reduction of 22% in CRC incidence rates and 33% in CRC mortality rates by 2030, which translates to approximately 280,000 averted new cases and 200,000 averted deaths from 2013 through 2030.

The 20% increase in screening uptake from 60% to 80% has a projected high impact on CRC mortality (33% reduction). This 33% matches well with our recent estimate that the majority (60%) of current CRC mortality is attributable to nonuse of screening.[209] The increase in screening uptake from 60% to 80% decreases the number of underscreened people by half and consequently reduced overall CRC mortality by roughly half the 'population attributable fraction'.

Within the underscreened population, the impact of '80% by 2018' will be larger than the 20-30% overall reductions in incidence and mortality for the population, because the majority of avoidable cases and deaths occur within the 40% of the

population that is underscreened. Underscreened individuals tend to have lower educational levels and income and lack health insurance.[89] Thus, a desirable effect of achieving '80% by 2018' is the potential to reduce CRC health disparities in the United States – an important HealthyPeople 2020 objective.[210]

To our knowledge, no prior study estimated the public health benefits of '80% by 2018'. Several studies have estimated the potential contribution of screening to decreases in CRC incidence and mortality in the United States.[198,199,211,212] Our estimates of screening benefits appear to be somewhat smaller than those from Ladabaum and Song[211] and larger than those from Edwards et al.[198] and Yang et al.[199] This is likely due to different study designs or periods, and differences in assumptions regarding the effectiveness of colonoscopy screening. For colonoscopy, the effectiveness of screening is less well-established than for other recommended screening tests due to the absence of evidence from randomized controlled trials. The effectiveness of endoscopy screening in the MISCAN model was recently increased based on the outcomes of the UK flexible sigmoidoscopy study.[117] This change explains the slightly higher impact of increasing screening uptake in this study compared with earlier studies.[198,212] We evaluated a more conservative assumption for colonoscopy efficacy where colonoscopy sensitivity was decreased by 50%; this decreased the impact of '80% by 2018' on incidence, but did not substantively influence the mortality benefits (data not shown).

There are some limitations to this study. First, we evaluated only one of two possible ways to increase screening rates in the United States, namely by expanding screening to previously unscreened people. An alternative way is to reduce the number of people who have been screened but not according to screening recommendations. In the latest NHIS from 2013, the proportion of the population which ever received a CRC test, but not within the recommended intervals, was 7.4%.[200] Thus, in a strategy of encouraging both higher guideline adherence in previously screened people and the participation of previously unscreened people, the former approach could contribute one third (7.4%) to the overall targeted increase of 22% in screening rates. This may lead to a somewhat lower public health impact of '80% by 2018' than we found, because the impact of screening is lower in previously screened people compared to unscreened people.

Second, we assumed that the proportion of endoscopy versus fecal-based exams and its quality remained the same in the population when increasing screening uptake, while higher uptake of FIT or other stool-based tests may be needed to achieve the ideal of 80% screening.[83,85] A higher proportion of stool-based tests than modeled may affect the projected benefits of increased screening uptake, although modeling analyses show that the potential benefit of 10-yearly colonoscopy and annual FIT may be comparable.[204] Colonoscopy quality is known to vary widely among

providers and is highly correlated with disease outcomes.[124] If expanding screening, in part, was achieved through examiners with lower detection rates then the benefits may be less than projected.

Finally, there may be CRC disparities between screened and underscreened populations beyond those attributable to screening.[213] If the background CRC risk in the underscreened population is higher and/or CRC survival poorer, the impact of reaching 80% screening by 2018 may be even larger.

The outcomes of this study were confined to CRC incidence and mortality in the population, and did not include years of life lost to CRC, costs and potential harms of screening. Previous analyses have indicated that CRC screening is likely highly cost-effective,[214,215] and may even be cost-saving,[28] making increasing screening not only desirable from a cancer-control perspective but also from a financial perspective. However, these analyses usually do not consider potential overuse of screening and surveillance,[62] program costs,[100] and especially, resources needed to bring in the people to reach 80% uptake of screening.

There are many barriers to increasing CRC screening uptake in the U.S., only some of which are the target of health care reforms under the Affordable Care Act .[82,216] Substantial coordinated effort is needed to achieve the goal of 80% CRC screening by 2018 goal in the United States. The results of our study indicate that such investments may be well-rewarded with long term reductions in CRC incidence and mortality of 22% and 33%, respectively, and the avoidance of 280,000 new CRC cases and 200,000 CRC deaths in less than two decades.

Chapter 4

Colorectal cancer screening: estimated future colonoscopy need and current volume and capacity

Article published online (AltMetric = 59):

DA Joseph, RG Meester, AG Zauber et al.
Cancer. 2016 Aug 15;122(16):2479-86.

ABSTRACT

BACKGROUND: In 2014 a national campaign was launched to increase colorectal cancer (CRC) screening rates in the U.S. to 80% by 2018; it is unknown if there is sufficient colonoscopy capacity to reach this goal. We estimate the number of colonoscopies needed to screen 80% of the eligible population with fecal immunochemical testing (FIT) or colonoscopy, and if there is sufficient colonoscopy capacity to meet the need.

METHODS: The Microsimulation Screening Analysis-colon (MISCAN-colon) model was used to simulate CRC screening test use in the U.S. (2014-2040), assuming the implementation of a national screening program in 2014 with FIT or colonoscopy with 80% participation. The 2012 Survey of Endoscopic Capacity (SECAP) estimated the number of colonoscopies that were performed and the number that could be performed.

RESULTS: If a national screening program started in 2014, by 2024, approximately 47 million FITs and 5.1 million colonoscopies would be needed annually to screen the eligible population with a program using FIT as the primary screening test; approximately 11 to 13 million colonoscopies would be needed annually to screen the eligible population with a colonoscopy only screening program. Based on the SECAP survey, an estimated 15 million colonoscopies were performed in 2012 and an additional 10.5 million colonoscopies could be performed.

CONCLUSIONS: The estimated colonoscopy capacity is sufficient to screen 80% of the eligible U.S. population with FIT, colonoscopy, or a mix of tests. Future analyses should take into account the geographic distribution of colonoscopy capacity.

INTRODUCTION

Although screening for colorectal cancer has been shown to effectively reduce the incidence of and mortality from the disease, only 58% of adults aged 50-75 years were up-to-date with CRC screening in 2013.[217] A recent initiative from the National Colorectal Cancer Roundtable (NCCRT), a coalition of public, private, and voluntary organizations, aims to increase CRC screening prevalence to 80% in the eligible population by 2018. Recent analyses estimated that reaching this goal would avert 280,000 new cases of and 200,000 deaths from CRC by 2030 and that 24.4 million people would need to be screened.[218,219] No studies have estimated the number of CRC screening tests that would need to be performed each year if 80% prevalence is achieved, and whether current colonoscopy capacity would meet increased demand. Over the past decade, colonoscopy use has increased rapidly and has become the most commonly used test to screen for CRC, while relative use of fecal occult blood testing has declined.

We used microsimulation modeling to estimate the expected number of colonoscopies to screen 80% of the eligible population with either fecal immunochemical tests (FIT) or colonoscopy over 10 years. We also conducted a national Survey of Endoscopic Capacity (SECAP) to estimate the number of colonoscopies performed in a year in the U.S., and the number of additional colonoscopies that could be performed (capacity). Resources, or capacity, are defined as non-monetary resources, such as number of staff, facility space, equipment and time needed to perform colonoscopies, and does not include the actual cost of the procedures paid for by individuals or insurers.

METHODS

Estimation of Screening Test Need

The MISCAN-colon model (**Chapter 2**) was used to simulate CRC screening test use in the U.S. (from 2014 to 2040), assuming the implementation of a nationwide screening program in 2014. The main outcome of the model was the number of colonoscopies required per year to screen 80% of the population. Screening was implemented over 10 years using FIT or colonoscopy as the primary screening test.

Simulated Scenarios

Age-specific use rates of colonoscopy, flexible sigmoidoscopy, and fecal occult blood test (FOBT) until the start of a hypothetical national screening program in 2014 were based on National Health Interview Survey (NHIS) data from 1987 through 2010.[217]

Based on these data, it was estimated that in 2013, 67% of U.S. adults aged ≥50 years had ever been screened with any test, 8.8% had a home FOBT within the last year, 4% had a sigmoidoscopy within the last 5 years, and 55% had a colonoscopy within the past 10 years. We assumed that there was no further increase in overall screening uptake in the period from the last NHIS to the start of the hypothetical screening program.

The model enrolled all U.S. adults aged 50 to 75 years into a national screening program over 10 years, starting with the first cohort in 2014, consisting of 1/10 of the age-eligible population. The model assumed that the remaining eligible population would continue to be screened at a projected estimate based on 2010 NHIS data until enrollment into the hypothetical national program. People were not invited for screening in the program until 1 year after their last FOBT, 5 years after their most recent sigmoidoscopy, or 10 years after their most recent colonoscopy. In the first scenario, we evaluated a program of annual FIT in which 80% of eligible adults participated; in the second scenario, we evaluated a program colonoscopy every 10 years with 80% participation. People with a positive FIT were referred for follow-up colonoscopy and people with an adenoma detected were followed with colonoscopy surveillance, with the interval (3 to 5 years) dependent on the number and size of adenomas detected on the most recent colonoscopy.[139,220]

Sensitivity analysis

In a sensitivity analysis we evaluated various alternative modeling scenarios, to inform readers on implications of other possible screening tests and adherence rates. Alternative modeling scenarios included: alternative primary screening tests, including annual guaiac fecal occult blood testing (gFOBT), 10-yearly computed tomographic colonography (CTC), and 5-yearly flexible sigmoidoscopy (FSIG); higher assumed participation rates (100%) for FIT and colonoscopy screening; and a scenario of currently observed test use patterns in NHIS (with both under- and over-use), with an assumed linear increase in overall screening participation rates from 58% in 2013 to 80% by 2018. Test performance characteristics used in the primary and sensitivity analysis are provided in Supporting Information, **Supplementary Table 4.1**.

Estimation of Endoscopic Capacity

The Survey of Endoscopic Capacity II (SECAP II) was conducted in 2012. The survey methodology was unchanged from the original SECAP study; a detailed description of the survey methodology has been published previously.[221] In brief, a list of all U.S. medical facilities known to have purchased or leased lower endoscopic equipment between January 1, 2006 and December 31, 2010 was obtained from three major endoscopic equipment manufacturers: Olympus America, Inc., Fujinon, Inc., and Pentax Precision Instrument Corporation. The lists were merged and duplicates removed to create a sampling frame. A random sample of 2100 facilities (31% of all facilities), stratified by region and location (urban or rural), was selected to participate in the survey. A telephone screening questionnaire was administered to confirm study eligibility and to identify the person in charge of endoscopy. Of the 2100 facilities, 258 (12%) were found to be ineligible (did not currently perform screening sigmoidoscopy or colonoscopy on adults or could not be located). A self-administered questionnaire, personalized cover letter, postage-paid return envelope, and $40.00 incentive were sent by Federal Express to a person identified by each eligible facility. Respondents were asked to provide an estimate of the total number of sigmoidoscopies and colonoscopies performed by all endoscopists at the practice site per week, the percentage of procedures performed by endoscopist specialty, and the additional number of sigmoidoscopies and colonoscopies that could be performed with no other investment of resources.

Of the 1842 eligible facilities, 1269 returned valid surveys (overall response rate 68.9%). To provide national capacity estimates, the universe of facilities was adjusted based on the ineligibility rate, and survey data were weighted to adjust for the sampling weight and non-response. Annual estimates of capacity were obtained by multiplying the weighted weekly estimates of current and potential capacity by the number of workweeks per year (50 weeks). Survey data were analyzed with Stata 12.1.

For the estimation of endoscopic capacity, two questions were critical to the analysis: 1) the number of procedures currently performed and 2) the additional number of procedures that could be performed. If answers to both of these questions were missing, the survey was excluded from analysis. If the survey was missing data for one of the two key question, then these values were imputed using a variation of the hot-deck method, as described previously.[221]

RESULTS

Simulation Results

Based on recent CRC screening patterns, an estimated 8.4 million FOBTs and 14 million colonoscopies were performed in 2013. Of these, approximately 3.3 million colonoscopies were estimated to have been performed for diagnostic and surveillance purposes (**Figure 4.1a**).

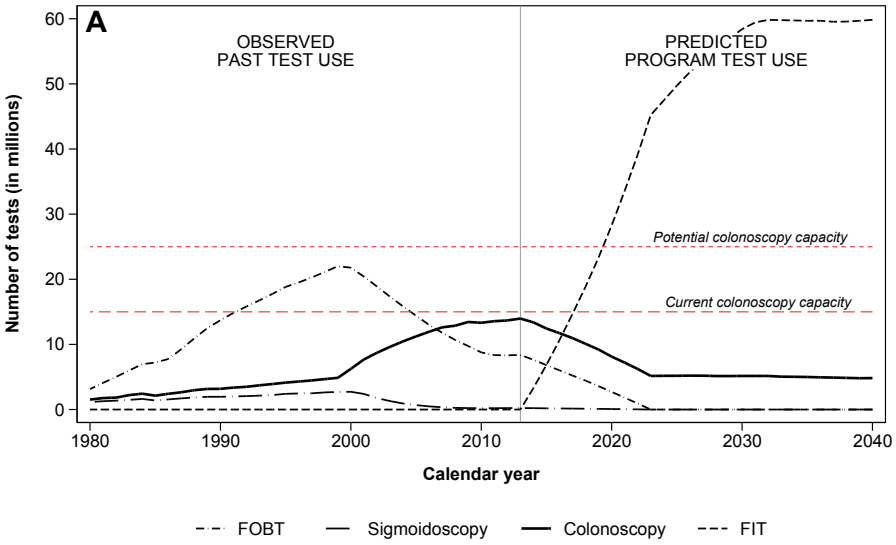


Figure 4.1a Number of colorectal cancer tests per year before and after start of hypothetical national screening program with FIT* in 2014, by test type
Abbreviations: FOBT = fecal occult blood test; FIT = fecal immunochemical test.

FIT Scenario

Assuming the introduction of a FIT screening program in 2014, a total of 3.3 million FITs would need to be performed to screen 80% of eligible adults aged 50 to 75 invited to the first round of screening (1/10 of the eligible population). The total number of colonoscopies needed in 2014 would be 13.4 million: 3.5 million for diagnostic or surveillance purposes, and 9.9 million for screening performed outside the program (**Figure 4.1b**). By 2024, approximately 47 million FITs and 5.1 million diagnostic (32%) and surveillance (68%) colonoscopies would have to be performed. The number of FITs would gradually increase to approximately 60 million tests annually by 2030, but the number of colonoscopies would remain steady.

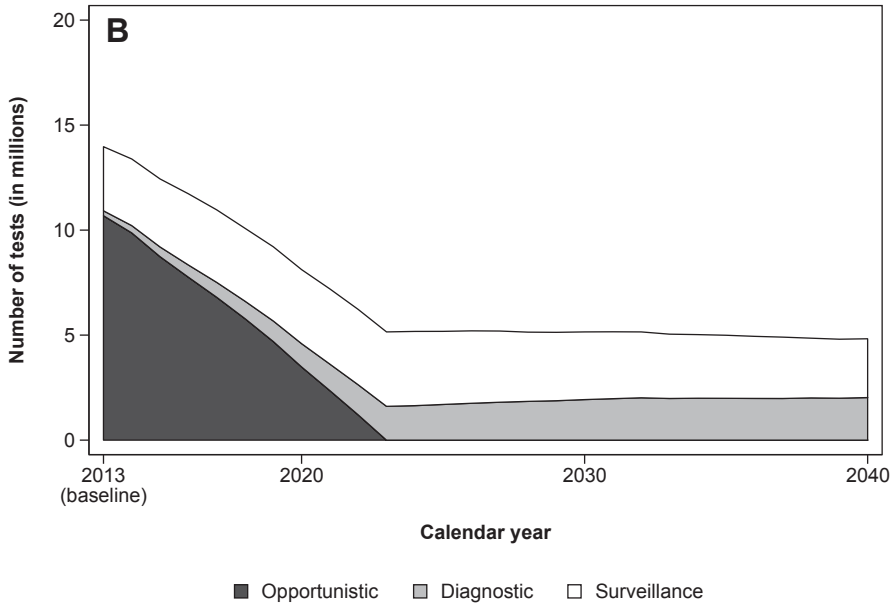


Figure 4.1b Number of colonoscopies per year, before and after start of a hypothetical national screening program with FIT, by colonoscopy indication

Colonoscopy Scenario

The introduction of a colonoscopy screening program in 2014 would require 12.8 million screening colonoscopies and 3.4 million diagnostic and surveillance colonoscopies (**Figures 4.2a-b**). By 2024, 11 to 13 million colonoscopies would have to be performed annually, with ~57% being performed for screening and~43% for surveillance, and remain level through 2030.

Sensitivity analysis

Estimated colonoscopy requirements assuming 80% participation of all eligible adults were similar for annual FIT, annual gFOBT, and 5-yearly CTC testing (**Figure 4.3**). FSIG every 5 years would require 16.3 colonoscopies in 2014, and 18-19 million sigmoidoscopies and colonoscopies annually by 2030. Continuation of currently observed test use patterns with 80% participation would require 23 million colonoscopies annually by 2030.

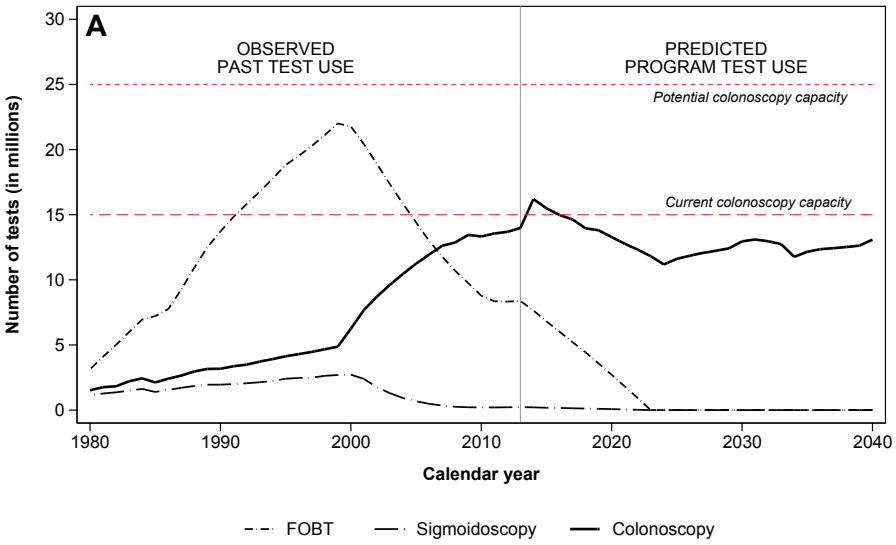


Figure 4.2a Number of colorectal cancer tests per year before and after start of a hypothetical national screening program with colonoscopy, by test type
Abbreviations: FOBT = fecal occult blood test.

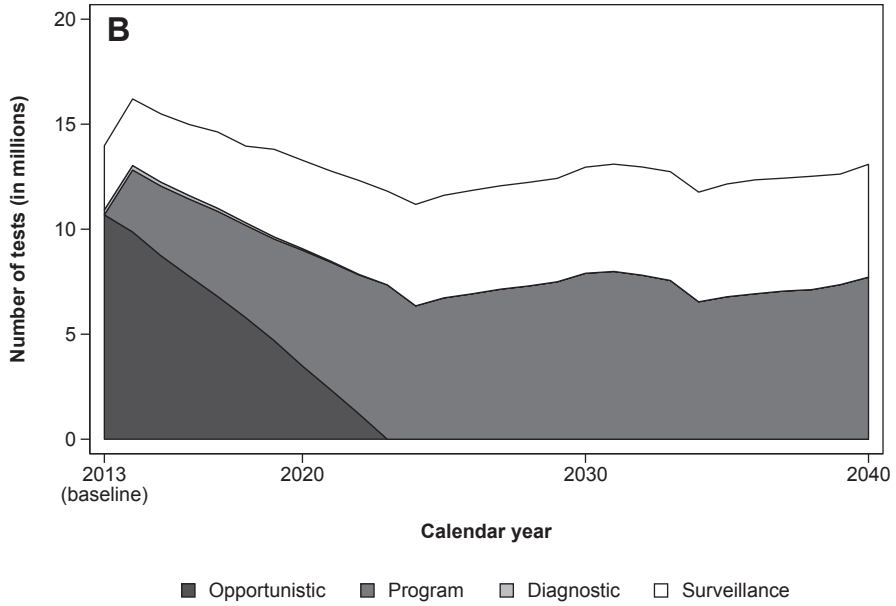


Figure 4.2b Number of colonoscopies per year, before and after start of a hypothetical national screening program with colonoscopy, by colonoscopy indication

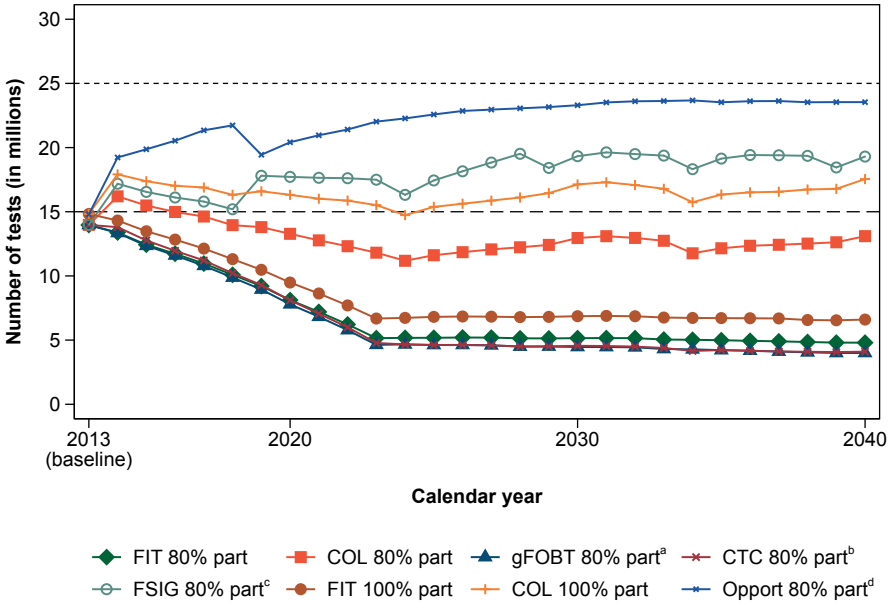


Figure 4.3 Predicted colonoscopy use under various modeling scenarios (in millions).
FIT = Fecal Immunochemical Test; COL = Colonoscopy; gFOBT = guaiac Fecal Occult Blood Test; FSIG = Flexible Sigmoidoscopy; Opport. = Opportunistic; part = participation.
[a] Annual testing with Hemoccult II.
[b] 10-yearly testing with CT colonography.
[c] 5-yearly testing with flexible sigmoidoscopy. Numbers represent sigmoidoscopy and colonoscopy use.
[d] In the scenario with opportunistic screening, we assumed future screening patterns according to age and type of test were similar to those observed in 2013 National Health Interview survey data. The screen rate was increased linearly from approximately 60% to 80% from 2013 to late 2018.

The required colonoscopy capacity with either FIT or colonoscopy screening with 100% participation was approximately one-third higher than the capacity needed for the base-case of 80% assumed participation. The 100% FIT scenario would require 14.3 million colonoscopies in 2014, and would require 68 million FITs and 6.9 million colonoscopies annually by 2030. The 100% colonoscopy scenario would require 14.1 million colonoscopies in 2014 and 17 million colonoscopies annually by 2030.

Survey

Of the 1269 facilities included in the final analysis, 767 (60.9%) were hospital departments, 403 (31.8%) were ambulatory endoscopy or surgery centers, 98 (7.7%) were physician practices and 1 was unknown (data not shown). The majority of survey respondents identified themselves as nurse administrators/managers (60.2%). The majority of sites were classified as urban (68.2%). After weighting, there were an

estimated 5988 (95% confidence intervals [CI] 5832 to 6144) facilities in the U.S. that performed any lower endoscopy in 2012. Of these, 5858 (97.8%) facilities reported performing colonoscopy and 1831 (30.6%) reported performing sigmoidoscopy.

Survey respondents estimated that 51.1 (95% CI 46.1 to 56.1) colonoscopies were performed per week (**Table 4.1**). Respondents estimated that 43.2% of colonoscopies were performed for screening. The total mean potential maximum number of colonoscopies that could be performed per week was 87.

Table 4.1 Current and potential number of colonoscopies, Survey of Endoscopic Capacity II - 2012

	Total (SE) [a]
Number of facilities [b]	5858 (202.6)
Current weekly number (mean)	51.1 (2.5)
Percent performed for screening (in millions)	43.2 (0.6)
Potential Maximum weekly number (mean)	87.0 (5.4)
Current annual[c] volume (in millions)	15.0 (1.2)
Potential annual[c] volume (in millions)	25.5 (2.4)
Available annual[c] capacity (in millions)	10.5 (2.6)

[a] SE = Standard error

[b] Facilities included hospitals, ambulatory surgery centers, and physician offices where colonoscopies were performed for the purpose of colorectal cancer screening of adults.

[c] Assuming 50 work weeks per year

Survey responses were weighted to determine national estimates for current and potential capacity to provide colonoscopies in the U.S. In 2012, approximately 15 million total colonoscopies were performed (**Table 4.1**). Respondents reported they could increase their colonoscopy volume to 25.5 million annually for an available capacity of 10.5 million colonoscopies annually.

DISCUSSION

This report estimates the number of colonoscopies that would be needed to screen 80% of the eligible population and compares this need to estimates of colonoscopy capacity. The MISCAN microsimulation model estimated that 13.4 million colonoscopies would be needed in the first year of a population CRC screening program with FIT, gradually declining to 5.2 million colonoscopies with full implementation of the program after 10 years. A colonoscopy program implemented over 10 years would require 16.2 million colonoscopies in the first year, and 12 to 13 million colonoscopies annually with full implementation. According to the survey, in 2012, 15 million colonoscopies were performed, of which respondents estimated that

42.3% (6.3 million) were performed for screening. Respondents indicated that an additional 10.5 million colonoscopies could be performed per year, suggesting that the increased demand for screening colonoscopy could be absorbed. The FIT screening program would require no screening colonoscopies, and the demand for diagnostic and surveillance colonoscopies could presumably be met by shifting currently available resources. The colonoscopy screening program would require approximately 7 million screening colonoscopies and 5 million surveillance colonoscopies annually. Assuming no change in available capacity, the increased demand for screening and surveillance colonoscopies is matched by currently available colonoscopy capacity as reported. Given that a colonoscopy screening program is, for a given participation level, the strategy with the highest colonoscopy demand, there would also be sufficient capacity to meet colonoscopy demand for most of the scenarios modeled in the sensitivity analysis (FIT only or colonoscopy only with 100% participation, and annual gFOBT, 5-yearly CTC or FSIG with 80% participation). If recently observed CRC test use patterns continued with 80% participation, estimated capacity could meet colonoscopy need within the estimated standard error (22-24 million needed annually vs. 23.1-27.9 estimated annual capacity).

The percentage of the adult population that is up-to-date with CRC screening has steadily increased over the past decade, primarily through increased use of colonoscopy.[89,222,223] Our data do not show a concomitant increase in the number of colonoscopies performed annually. Although the SECAP survey is cross-sectional, and may not have captured a true rise and subsequent decline in the number of colonoscopies performed, at least one other study found that the use of screening colonoscopy increased prior to the recent economic recession, then subsequently declined.[224] After rapid growth from 2000 to 2006, a decline in the number of colonoscopies performed per Medicare beneficiary has also been noted.[225]

The MISCAN microsimulation model estimated that approximately 13.7 million colonoscopies were performed in 2012 for screening and follow-up; analysis of the 2012 Behavioral Risk Factor Surveillance System (BRFSS)[226] and of the 2010 NHIS[217] estimated that 14.9 million people and 11.4 million people respectively had a colonoscopy within the past year. The 2012 SECAP estimate of 15 million colonoscopies performed closely matches these estimates. Of note, the number of adults aged 50 years or older that reported sigmoidoscopy or colonoscopy within the previous year remained largely unchanged from the 2003 (11.3 million) to the 2010 (11.8 million) NHIS, despite a substantial increase in the proportion of adults in this age group that reported being up-to-date with CRC screening by colonoscopy within 10 years.[217]

In the base-case analysis, future test use for a national CRC screening program with either FIT or colonoscopy estimated that 5 to 16 million colonoscopies would be needed annually, assuming 80% of the eligible population would participate.

These model projections included only colonoscopies performed for screening, diagnostic, or surveillance purposes and assumed no under- or overuse of screening and therefore may have under-estimated actual test need in these scenarios. The 2012 SECAP survey estimated that 43.2% of colonoscopies were performed for screening purposes consistent with previous estimates of 38% to 49.7%.[224,227] Surveillance colonoscopies have accounted for up to one-quarter of colonoscopies performed, suggesting that a substantial proportion of colonoscopies performed are for reasons other than screening or surveillance (i.e., diagnostic).[227,228] Several studies of the Medicare population have found over- and under-use of both screening and surveillance colonoscopies.[62,63,229-233] In our sensitivity analysis, continuation of current CRC test use patterns, reflecting current patterns of under- and overuse, required substantially more colonoscopies than even the colonoscopy only scenario. The modeled colonoscopy need may also have been overestimated as we assumed 80% adherence to screening and surveillance. Despite concerted efforts to increase CRC screening rates in the population, rates remain well below 80% due to a variety of patient, provider, and system level factors.[222,234,235] Among those who are screened for CRC with FIT or other tests that require follow-up with colonoscopy, adherence to follow-up is well below 100%.[236,237]

Full implementation of a national CRC screening program with FIT would require approximately 5 million diagnostic and surveillance colonoscopies annually. While the number of colonoscopies required is practically achievable, a national FIT program would also require nearly 60 million FITs annually by 2040. FIT does not require many resources on the part of the patient (can be done at home, does not require bowel preparation or dietary changes), but a complete FIT screening program can require substantial additional resources to ensure that test kits are distributed to the eligible population, remind people to complete and return the kits, ensure complete follow-up for those with positive results, process all returned kits in the provider's office or in the lab, and to ensure that all eligible adults repeat the test annually.[238,239] It is unknown if adequate resources exist to implement a FIT screening program on such a large scale. Our model estimates that a national colonoscopy screening program would require substantially more colonoscopies annually than a FIT program. It is unknown if it is feasible to shift resources towards more screening and surveillance and if sufficient capacity would remain to perform colonoscopies needed for other reasons.

This study is subject to some other limitations. First, our model estimates of future test need for a national CRC screening program were based on recent population projections, currently available screening methods, and current screening guidelines which may not apply to the entire time horizon of the study. Second, we could not validate directly the number of colonoscopies that survey respondents indicated they

were performing or could perform. Respondents were asked to estimate the number of additional colonoscopies they could do without additional resources, but it is unknown if the estimate truly reflects what could be done without changes to current practice or if it reflects shifting resources away from other procedures. Analysis of additional SECAP questions indicated that there were limiting factors to increasing capacity and, if needed, facilities would invest in additional resources (physicians, nurses, equipment, etc.) to increase capacity (Supporting Information, **Supplementary Tables 4.2-3**). As described earlier, our estimate of annual colonoscopy volume was consistent with estimates from other sources.[217,226] Third, the survey sampling frame included facilities that purchased or leased equipment between 2006 and 2010. This excludes facilities that use equipment purchased or leased outside of this time frame and may underestimate the number of colonoscopies currently performed and available capacity. Fourth, the study was not designed to model market forces as it relates to the supply of colonoscopy in response to increasing demand (e.g., the market could respond by increasing the supply of endoscopists). Fifth, this study could not account for the geographic distribution of CRC screening need or of colonoscopy capacity. The survey was not designed to estimate colonoscopy capacity at the local level, and simulating future screening need at this level would require an impractical number of models (to account for population size and past screening behavior for each geographical unit).

CRC screening is conducted with a variety of tests, most commonly colonoscopy and less frequently with FOBT or FIT. While it is unlikely that all eligible adults will be screened with a single test type, this analysis shows that the estimated colonoscopy capacity would be sufficient to screen with a mix of tests. Future analyses should take into account the geographic distribution of colonoscopy capacity and screening need, to determine if there is a surplus of capacity in some areas of the country and insufficient capacity in others.

APPENDIX 4

Supplementary Table 4.1 Colorectal cancer screening test performance assumptions, Microsimulation Screening Analysis-Colon (MISCAN-colon)

Performance characteristic	Test				
	Optical Colonoscopy	Flexible Sigmoidoscopy	CT Colonography	FIT (OC-Sensor, >20 ng/g cutoff)	FOBT (Hemoccult II)
Sensitivity per lesion					
Adenomas ≤ 5 mm	75%	75%	0.0%	0.0%	0.0%
Adenomas 6 - 9 mm	85%	85%	75.7%	4.4%	1.3%
Adenomas ≥ 10 mm	95%	95%	85.9%	13.1%	6.5%
Early stage I-IV cancer [a]	95%	95%	95%	52%	18.2%
Late stage I-IV cancer [a]	95%	95%	95%	83.5%	50.8%
Specificity [b]	100%	100%	91.4%	97.6%	98%
Completeness [c]	98%	-	-	-	-

FIT = Fecal Immunochemical Test; FOBT = Fecal Occult Blood Test.

[a] We assumed that fecal testing is more sensitive in preclinical cancers that are close time-wise to becoming symptomatic, i.e. towards the end of the occult invasive period. This assumption showed good concordance with guaiac fecal occult blood test trial results.

[b] The probability of a false positive result was random in the base-case analysis, and independent of person or lesion. We assumed perfect specificity for colonoscopy and sigmoidoscopy with pathological follow-up examination.

[c] This is the proportion of colonoscopies visualizing the maximum point of reach of the endoscope, i.e. the cecum. Sigmoidoscopy was assumed to reach the splenic flexure in 80% of examinations.

Supplementary Table 4.2 Primary limiting factor to performing additional colonoscopies [a], Survey of Endoscopic Capacity II – 2012.

Primary Limiting Factor	Percentage of Facilities[b] (SE[c])
Insufficient time (few open appointments)	6.5% (0.6)
Insufficient utilization due to cancellations (“no shows”)	5.7% (0.6)
Insufficient number of physicians available to perform procedures	34.0% (1.2)
Insufficient nursing staff to assist with procedures	8.9% (0.7)
Insufficient ancillary staff to help with room turnover	0.6% (0.2)
Insufficient staff or physicians to monitor sedation or anesthesia	2.1% (0.4)
Insufficient procedure rooms	6.9% (0.7)
Insufficient preparatory and/or recovery areas	8.2% (0.7)
Insufficient endoscopes or monitors	3.6% (0.5)
Insufficient reimbursement	4.0% (0.5)
Other [d]	19.5% (1.0)

[a] Respondents were asked “What are the limiting factors to performing more colonoscopies at this practice site?”, then were asked “What is the primary limiting factor?” Percentages reflect the proportion of respondents that chose the option as the primary limiting factor.
[b] Facilities included hospitals, ambulatory surgery centers, and physician offices where colonoscopies were performed for the purpose of colorectal cancer screening of adults.
[c] SE=Standard error.
[d] Most commons responses to other were competition from other facilities and lack of patient referrals.

Supplementary Table 4.3 Measures to address increased need for colonoscopy [a], Survey of Endoscopic Capacity – 2012

Measure	Percentage of Facilities[b] (SE[c])
Increase proportion of work day allotted to procedures	59.3% (1.3)
Modify block scheduling	56.9% (1.3)
Use patient navigators or reminder calls to decrease “no shows” or cancellations	37.3% (1.2)
Increase physician staff	55.3% (1.3)
Increase/hire non-physician endoscopists to do procedures	5.4% (0.6)
Increase nursing staff to assist with procedures	68.1% (1.2)
Increase ancillary staff to help with room turnover	51.6% (1.3)
Increase staff or physicians to help monitor sedation/anesthesia	41.2% (1.3)
Establish a larger screening unit/more procedure rooms	36.5% (1.2)
Establish additional preparatory and/or recovery areas	39.9% (1.3)
Purchase or lease more equipment	52.0% (1.3)
Other	4.5% (0.5)
Not applicable, not planning to perform more procedures	8.6% (0.7)

[a] In response to the question “If the demand for colonoscopies were to exceed this practice site’s current capacity to perform colonoscopies, what steps would this practice site take to meet that increased demand?” Respondents could select all options that applied.
[b] Facilities included hospitals, ambulatory surgery centers, and physician offices where colonoscopies were performed for the purpose of colorectal cancer screening of adults.
[c] SE=Standard error.

Chapter 5

Colorectal cancer deaths attributable to nonuse of screening in the United States

Article published (AltMetric = 15):

RG Meester, CA Doubeni, I Lansdorp-Vogelaar et al.
Ann Epidemiol. 2015 Mar;25(3):208-213.e1.

ABSTRACT

PURPOSE: Screening is a major contributor to colorectal cancer (CRC) mortality reductions in the U.S., but is underutilized. We estimated the fraction of CRC deaths attributable to nonuse of screening to demonstrate the potential benefits from targeted interventions.

METHODS: The established MISCAN-colon microsimulation model was used to estimate the population attributable fraction (PAF) in people aged ≥50 years. The model incorporates long-term patterns and effects of screening by age and type of screening test. PAF for 2010 was estimated using currently available data on screening uptake; PAF was also projected assuming constant future screening rates to incorporate lagged effects from past increases in screening uptake. We also computed PAF using Levin's formula to gauge how this simpler approach differs from the model-based approach.

RESULTS: There were an estimated 51,500 CRC deaths in 2010, about 63% (N~32,200) of which were attributable to non-screening. The PAF decreases slightly to 58% in 2020. Levin's approach yielded a considerably more conservative PAF of 46% (N~23,600) for 2010.

CONCLUSIONS – The majority of current U.S. CRC deaths are attributable to non-screening. This underscores the potential benefits of increasing screening uptake in the population. Traditional methods of estimating PAF underestimated screening effects compared with model-based approaches.

INTRODUCTION

Both the absolute number of cases as well as the incidence and disease-related mortality rates for colorectal cancer (CRC) have declined over the last three decades despite a high prevalence of risk factors, in contrast to trends observed in some other countries.[45] Evidence indicates that the increasing use of CRC screening has been the major contributor to the declining incidence and mortality rates in the U.S. from this disease.[198,199] However, screening remains underutilized, suggesting that a substantial proportion of current CRC deaths in the U.S. are avoidable. This has galvanized public action on increasing the uptake of screening;[240] however, lack of clarity persists regarding the proportion of current CRC deaths occurring as a result of nonuse of screening, and thus the potential public health benefits from increasing screening uptake.

The population attributable fraction (PAF) proposed by Morton Levin in 1953 has been widely used to assess the proportion of a disease outcome that occurs as a result of exposure to a risk factor, and thus the potential benefits from public health interventions to eliminate that exposure.[241] This concept, which is a function of the level of exposure to the risk factor and the size of the effect of exposure on the disease outcome, has been previously applied to assess the impact of underutilization of CRC screening on disease mortality.[201] Using this approach, Stock and colleagues reported that about 28 – 44% of deaths from CRC in the U.S. in 2005 may be attributable to nonuse of colonoscopy. However, this study used somewhat conservative estimates for the effect of colonoscopy screening that may not be applicable for the U.S.[242-244] Also, the study did not consider specific features of CRC epidemiology that are important for valid estimation of PAF. First, apart from colonoscopy, flexible sigmoidoscopy or fecal occult blood tests are also used for screening in the U.S., and therefore need to be considered in estimating PAF. Second, CRC is a heterogeneous disease characterized by a long latency between risk factor exposure and outcome. Mortality benefits from screening are derived not only from cancer detection, but also from the detection and treatment of precursor or early more curable invasive lesions. Thus, valid estimates of PAF require the consideration of benefits of screening that are realized over long time periods after the test date. Finally, patterns of exposure to CRC screening have evolved since the 1980s. According to data from the National Health Interview Survey (NHIS), the proportion of the U.S. population recently exposed to CRC screening tests increased from about 39% in 2000 to 58% in 2010.[89,222]

In the present study, we used microsimulation modeling to estimate the PAF of U.S. CRC deaths from non-screening. We compared these PAFs with an estimate of

PAF using Levin's formula to gauge how this simpler more accessible approach may differ from the microsimulation approach.

METHODS

Population attributable fraction

The population attributable fraction (PAF) for CRC is defined as the proportion of CRC deaths in adults who are age 50 years or older that is due to non-receipt of screening as recommended by national guidelines. Analogous to the first definition discussed by Rockhill and colleagues, a short treatise on the most common definitions used for PAF, this is expressed algebraically as:[245]

$$PAF = \frac{R_T - R_0}{R_T} = \frac{RR_{T/0} - 1}{RR_{T/0}} \quad (1)$$

where R_T is the observed CRC mortality risk within the population per year, R_0 is the risk in those screened (unexposed) per year, and $RR_{T/0}$ is the ratio. We used the MISCAN-colon model (**Chapter 2**) to generate the entries R_T and R_0 in definition (1). To compare the model approach and simple approach, the risk in the absence of screening, R_1, was also assessed. Since the use of screening, disease incidence and mortality, and risk of death from competing causes change over a person's lifetime, we derived PAF according to three age strata (age 50–64, 65–74, 75 and older). It was first derived for calendar year 2010 based on observed patterns of exposure to non-screening from national survey data up to 2010, and then extended to 2030, assuming a constant rate of exposure to screening after 2010 to explore the lagged effects from recent increases in screening uptake. See the supplementary appendix for more precise definitions of PAF according to stratum and calendar year.

This study was conducted within the National Cancer Institute's (NCI) Cancer Research Network (CRN) and as part of the NCI-funded Population-Based Research Optimizing Screening through Personalized Regimens (PROSPR) consortium. The aim of PROSPR is to conduct multi-site, coordinated, trans-disciplinary research to evaluate and improve cancer screening processes.

MISCAN-colon microsimulation population

The MISCAN-colon microsimulation model was used to stochastically generate a virtual population similar to the U.S. population in terms of the life expectancy and the natural history and occurrence of CRC. This model was defined for the 1980 – 2030 period, to cover both historical and possible future patterns of screening use and the corresponding CRC mortality effects. U.S. birth and all-cause mortality for

the model were based on U.S. Census Bureau population estimates from 2000[205] and generational U.S. Berkeley Mortality tables,[206] respectively.

Exposure to non-screening

To derive PAF, we simulated two scenarios on the uptake of screening in the U.S. First, we closely replicated age- and test-specific screening patterns for the U.S. as observed in 8 waves of NHIS from 1987 – 2010 (**Figure 5.1**). The NHIS is a cross-sectional survey with a complex design on a nationally representative sample of the U.S. population.[200] Questions regarding the use of CRC screening tests were asked during the following survey years: 1987, 1992, 1998, 2000, 2003, 2005, 2008, and 2010. The estimated overall screening rate in 2010 (ages 50-100 years) was 59%. We assumed screening rates levelled off at ~60% (i.e. a 40% non-screening rate) after 2010. Screening as measured in the NHIS is comprised of home-based fecal occult blood testing, and endoscopy (particularly flexible sigmoidoscopy, or optical colonoscopy).

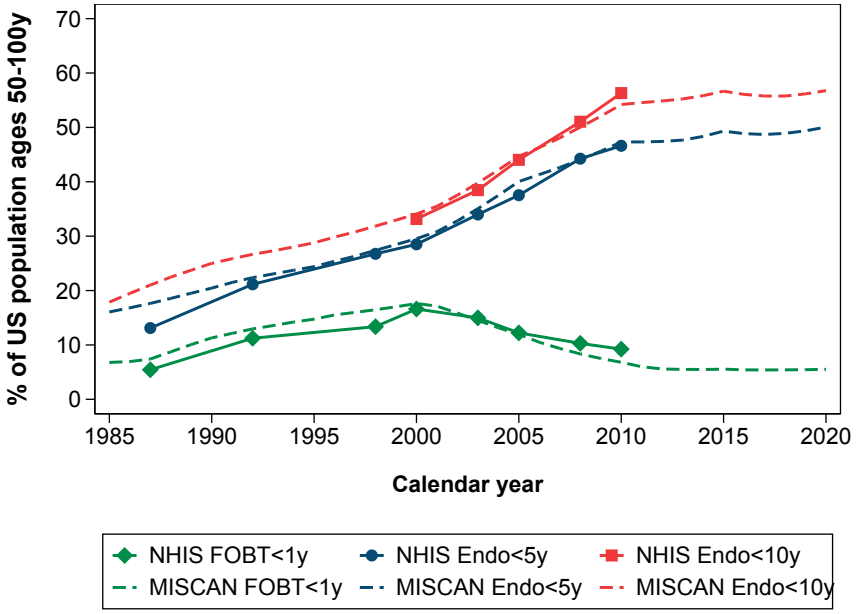


Figure 5.1 Colorectal cancer screening trends in National Health Interview Survey data and MISCAN. [a]

Abbrevations: NHIS = National Health Interview Survey; FOBT = fecal occult blood test; Endo = endoscopy.

[a] The red line plots the proportion of U.S. population which had a home FOBT in the previous year, the blue and green lines plot the proportions which had an endoscopy in the previous 5 or 10 years, respectively.

In the second scenario to assess the mortality risk from CRC that persisted despite complete screening of the population, after 1980 everyone was assumed to be fully compliant with using a single test (screening colonoscopy) at ages 50, 60 and 70 in accordance with U.S. guideline recommendations.[246]

In the first scenario above patients screened with a positive fecal test or sigmoidoscopy were invited for a diagnostic colonoscopy. The assumed adherence rate was 80%. In both of the above scenarios patients in whom precancerous adenomas were detected during colonoscopy were invited for surveillance colonoscopy at 3 – 5 yearly intervals in accordance with U.S. guideline recommendations for polyp size, number and histology.[139] The adherence rate for surveillance colonoscopy was 80% and 100%, respectively, for the two scenarios.

Screening and treatment effects

The effects of screening follow from the test performance assumptions in **Table 5.1**. We defined for each test the sensitivity and specificity for adenomas and adenocarcinomas, and in the case of endoscopic procedures, the extent of the colon evaluated by the exam. For detected incident adenomas, we assumed a 100% efficacy of treatment; for detected cancers, stage-specific survival was based on SEER mortality data for people with CRC diagnosed between 2000 – 2003. A model including these test characteristics was previously validated to data from trials on the effectiveness of sigmoidoscopy[247] and of fecal occult blood tests (**Chapter 2**).[157,188,189] The latter also included validation of the effect of colonoscopy after a positive test.

Table 5.1 Test performance assumptions in MISCAN

Performance characteristic	Colono-scopy [a]	Sigmoido-scopy [b]	FOBT [c]
Sensitivity:			
Adenomas ≤ 5 mm	0.75	0.75	-
Adenomas 6 - 9 mm	0.85	0.85	0.013
Adenomas ≥ 10 mm	0.95	0.95	0.065
Stage I adenocarcinoma	0.95	0.95	0.182 / 0.508
Stage II adenocarcinoma	0.95	0.95	0.182 / 0.508
Stage III adenocarcinoma	0.95	0.95	0.182 / 0.508
Stage IV adenocarcinoma	0.95	0.95	0.182 / 0.508
Specificity:	NA	NA	0.02
Reach endoscope:	Cecum	Splenic Flexure	NA
Completeness rate:[d]	0.98	0.8	NA

Abbreviations: FOBT = Fecal Occult Blood Testing; NA = Not Applicable

[a] Colonoscopy sensitivity for each adenoma, and completeness of colonoscopy were based on a systematic review of adenoma miss rates in tandem colonoscopy studies by Van Rijn and colleagues.[163]

[b] Sensitivity of sigmoidoscopy was also based on van Rijn and colleagues.[163]

[c] We assumed that fecal occult blood testing is more sensitive in preclinical cancers that are close time-wise to becoming symptomatic. This assumption showed good concordance with Fecal Occult Blood Test trial results.[190]

[d] This is the proportion of endoscopies visualizing the maximum point of reach of the endoscope.

Absolute CRC mortality risks

The CRC mortality risk was determined by the model assumptions for the risk of CRC, levels of screening uptake, and the effects of screening and treatment. The 2010 (baseline) mortality rate over all ages was aligned with the SEER mortality database by scaling the CRC incidence rate in the model (**Figure 5.2**).[248] Absolute mortality numbers for 2010 were derived by multiplying the mortality rates with 2010 population estimates from the U.S. Census bureau.[249]

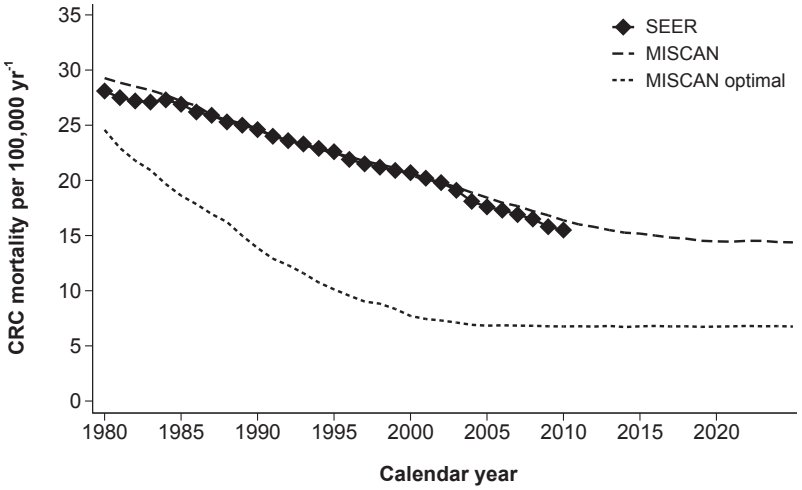


Figure 5.2 U.S. age-standardized[a] colorectal cancer (CRC) mortality rates by calendar year in Surveillance Epidemiology and End Results program (SEER) data and MISCAN
[a] Adjusted to the total 2000 U.S. standard population

Table 5.2a U.S. colorectal cancer deaths in 2010 attributable to nonuse of screening according to MISCAN

Variable	Population subgroup by age			
	50-64	65-74	75-100	All
Total population (million) [a]	59.1	21.9	18.6	99.6
Estimated number of CRC deaths without screening (MISCAN)	19,800	23,000	51,800	93,400
Actual number of CRC deaths in the population [b]	12,700	12,300	26,500	51,500
Estimated number of CRC deaths with full uptake of screening (MISCAN) [c]	7100	4200	7300	19,300
CRC deaths prevented by current screening (deaths if theoretical no screening – actual deaths)	7100	10,800	25,200	41,900
CRC deaths attributable to residual non-screening (actual deaths – deaths if 100% screening)	5600	8000	19,200	32,200
Attributable fractions:				
Fraction of CRC deaths attributable to non-screening if theoretical no screening, %	64%	82%	86%	79% [d]
Fraction of actual CRC deaths attributable to non-screening, %	44%	65%	72%	63%

Abbreviations: CRC = Colorectal cancer
[a] Population estimates were based on U.S. Census Bureau population estimates[249]. The overall population size in MISCAN was scaled to this number.
[b] CRC mortality numbers were derived by multiplying CRC mortality rates from 2010 SEER data with the population estimates from the U.S. Census Bureau.[248,249]
[c] This was defined as having screening colonoscopy at ages 50, 60 and 70 and lifetime surveillance follow-up of patients with adenomas detected in screening.
[d] Thus, the estimated overall relative risk for colonoscopy screening according to guideline recommendations was 0.21.

Alternative approach to assess PAF

We also derived PAF using the formula as proposed by Morton Levin in 1953, to help gauge the difference of this simpler more accessible approach with the model-based estimate.[241] Similar to the second definition in Rockhill and colleagues, this can be expressed algebraically as:[245]

$$PAF = \frac{P_1(RR_{1/0} - 1)}{P_1(RR_{1/0} - 1) + 1} \quad (2)$$

Here P_1 is the population proportion exposed to nonuse of screening, and the $RR_{1/0}$ is the ratio of the CRC mortality risks or rates in the non-screened versus the adequately screened population. This approximation is based on the assumption that the risk in the total population can be derived by linear interpolation of the risks in the non- and adequately screened groups ($R_T \sim P_1R_1 + (1 - P_1)\ R_0$), which is valid only under stringent conditions such as no confounding.[245] In this study, the parameters for equation (2) were derived from the same NHIS data used to inform the model on screening uptake in the U.S., and a large prospective cohort study for the effect of colonoscopy use.[166] Since the formula allows for a single parameter on screening uptake, we used the most recent (2010) NHIS wave to estimate the proportion exposed to non-screening (**Supplementary Table 5.1**). As risk ratio we used the age-adjusted hazard rate for colonoscopy use of 0.32 (95% CI: [0.24, 0.45]) derived by Nishihara and colleagues.[166] Again, more precise definitions according to age stratum and calendar year are provided in the supplementary appendix.

RESULTS

In 2010, the overall estimated number of CRC deaths in the U.S. was 51,500 (**Table 5.2a**). From this total, an estimated 12,700 occurred within the age stratum 50 – 64, 12,300 occurred within the age stratum 65 – 74, and 26,500 occurred within age stratum 75 and older.

In an ideal scenario of 100% uptake of screening (i.e. 100% uptake of 10-yearly colonoscopy screening), the microsimulation model estimated the expected number of CRC deaths to be 19,300 (**Table 5.2a**). This means that 32,200 CRC deaths out of the actual total of 51,500 in 2010 were attributable to nonuse of screening, which equates to a PAF of 63%. In analyses stratified according to age, the PAF was 44% for persons 50 – 64 years of age, but was 65% for those aged 65-74. On the assumption that screening rates remained at the 2010 level of ~60% into future years, the fraction of CRC deaths attributable to nonuse of screening decreased slightly over time to 58% in 2020 (**Figure 5.3**).

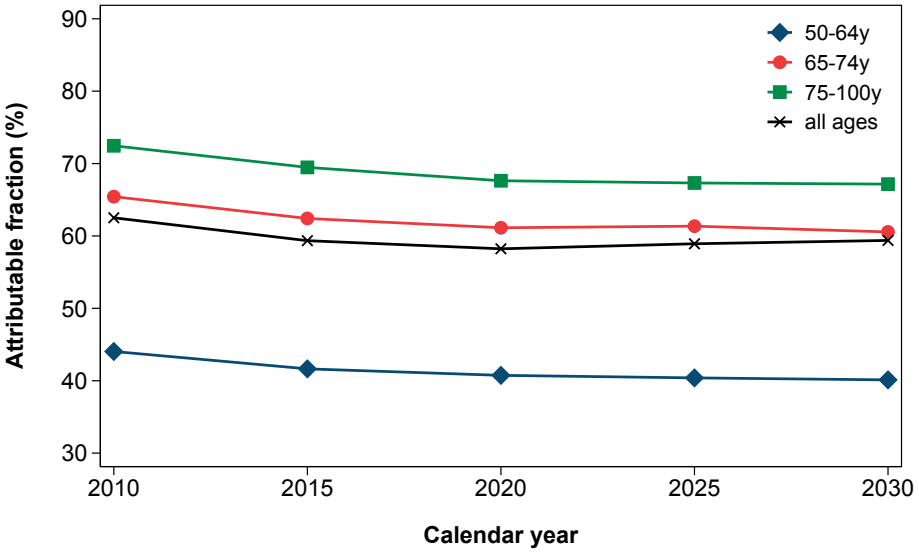


Figure 5.3 Projected CRC mortality fractions attributable to nonuse of screening [a]
[a] The mortality rates were not standardized for age; future estimates were based on a scenario of constant screen rates of ~60% after 2010

Levin's formula approach to estimate PAF yielded more conservative estimates of the fraction of CRC deaths attributable to underuse of screening. With this formula, 23,600 CRC deaths out of 51,500 in 2010 were attributable to underuse of CRC screening for a PAF of 46% (**Table 5.2b, Figure 5.4**). For the 50-64 year-old age group, the PAF was 49%, whereas for those 65-74 years old, the PAF was 41%, which was substantially lower than the result of the microsimulation approach.

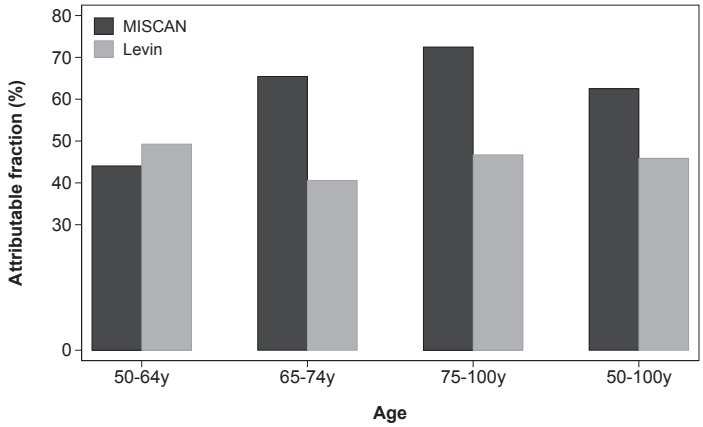


Figure 5.4 Proportion of U.S. colorectal cancer deaths in 2010 attributable to nonuse of screening by two approaches

Table 5.2b. U.S. colorectal cancer deaths in 2010 attributable to nonuse of screening according to Levin's formula

Variable	Population subgroup by age			
	50-64	65-74	75-100	All
Total population (million) [a]	59.1	21.9	18.6	99.6
Estimated number of CRC deaths without screening (MISCAN)	-	-	-	-
Actual number of CRC deaths in the population [b]	12,700	12,300	26,500	51,500
Estimated number of CRC deaths with full uptake of screening (MISCAN) [c]	6400	7300	14,100	27,900
CRC deaths prevented by current screening (deaths if theoretical no screening – actual deaths)	-	-	-	-
CRC deaths attributable to residual non-screening (actual deaths – deaths if 100% screening)	6200	5000	12,400	23,600
Attributable fractions:				
Fraction of CRC deaths attributable to non-screening if theoretical no screening, %	68%	68%	68%	68%
Fraction of actual CRC deaths attributable to non-screening, % [Min,Max] [d]	49% [36,59]	41% [28,50]	47% [34,57]	46% [33,56]

Abbreviations: CRC = Colorectal cancer

[a] Population estimates were based on U.S. Census Bureau population estimates[249]. The overall population size in MISCAN was scaled to this number.

[b] CRC mortality numbers were derived by multiplying CRC mortality rates from 2010 SEER data with the population estimates from the U.S. Census Bureau[248,249]. Likewise, numbers corresponding with the attributable fraction of CRC mortality were derived by multiplying the estimated PAF based on relative mortality rates with the observed number of deaths.

[c] Based on the age-adjusted hazard rate for colonoscopy use derived by Nishihara and colleagues [166]

[d] The minimum to maximum range was based on using respectively the 95% upper and lower confidence bound for the efficacy of screening reported by Nishihara and colleagues[166].

DISCUSSION

In the present study we used a Microsimulation Screening Analysis (MISCAN) model to assess the fraction of colorectal cancer (CRC) deaths in the U.S. population among people aged 50 or older that is attributable to nonuse of screening as recommended by U.S. national guidelines. Of the estimated 51,500 CRC deaths in 2010 in the U.S., about 63% (N~32,200) were attributable to non-screening. Under a scenario in which the screening rates attained in 2010 remained unchanged until 2030, the future population attributable fraction (PAF) attributable to nonuse of screening decreased to about 58% by 2020 due to the long-term cancer-preventive effects of adenoma removal after recent increases in screening uptake. Compared with the model-based approach, the traditional approach using the formula proposed by Levin, which utilizes static measures of screening and risk, resulted in a more conservative estimate of 46% (N~23,600) of CRC deaths in 2010 that were attributable to non-screening.

The PAF is an informative concept in providing public health policy makers with a ceiling for potential risk reductions achievable through interventions targeting the elimination of risk factors.[245] In this study we found that considerable reductions in CRC mortality of up to 63% are possible if the screening uptake in the U.S. is maximized (100% uptake). Unfortunately, the likelihood of this outcome occurring in the foreseeable future seems small. Healthcare accessibility is still a serious problem for roughly one quarter of the U.S. population,[70] and screening is underutilized particularly by populations with lower socioeconomic status, including the uninsured.[82] Even if recent health insurance reforms fulfill their promise of minimizing financial barriers to access for underserved populations,[250] the question remains whether uptake will get beyond a level of 80%. Integrated health care delivery systems have been successful in achieving compliance rates of around 80%, such as in the Kaiser Permanente Northern and Southern California member populations,[95] where all eligible adults not up-to-date with screening by endoscopic methods receive a fecal hemoglobin test over the mail (out-reach), and are reminded during primary care visits (in-reach).[239] This 80% screening rate has also been declared a national goal for the U.S. by 2018;[203] the observation that it has already been achieved in some large populations suggest this goal is, at the least, feasible. Assuming linearity in the effects of screening to screening uptake (the basic assumption behind the formula by Morton Levin[241]), our results indicate that a reduction of ~30% (half of the effect when attaining full compliance) in CRC mortality might be expected if the US succeeds in attaining the 80% screening rate.

The model-based approach resulted in a substantially higher PAF than the traditional approach using Levin's formula. This stemmed from a number of factors. First, the model incorporated long-term patterns and effects of screening in the

population, while the simpler approach used a static measure for the proportion exposure to nonuse of screening. Thus the traditional approach, for example, could not incorporate in its calculation of CRC mortality for 2010, the benefits from cancer prevention from the removal of adenomas provided by screening exams received many years earlier. Given the steep increase of screening rates in recent years, this may have contributed to the underestimation of the 2010 PAF with this method. The model suggests that this underestimation may have accounted for about one-quarter to one-third of the difference between the approaches, given the narrower gap between the two approaches for 2020.

The remainder of the difference in PAF between model-based and simple approach was attributable to different assumptions for the efficacy of screening. First, while in MISCAN the risk ratio over all ages corresponding with the use of colonoscopy screening according to guidelines was approximately 0.21 (**Table 5.2a**), a ratio of 0.32 for colonoscopy use in general was used in the simple approach. This larger risk reduction for screening leveraged the model-based PAF. Assuming a lower risk ratio of 0.21 the simple approach would have generated a PAF closer to the model estimate. Further, while the model allowed for disparities in effects of screening according to age and current versus optimal screening practice, a uniform risk ratio was applied in the simple approach. Because the model settings induced stronger effects of screening in the older age strata, the PAF difference was most pronounced in higher ages. The simpler (non-model based) approach in this study did lead to an overall PAF similar to the 44% found for 2005 in the previous study by Stock and colleagues,[201] who also utilized the simpler traditional approach to derive PAF.

We used NHIS data to inform this study on the current and past exposure to screening in the U.S., which is subject to potential biases. First, NHIS is a cross-sectional survey with repeated measurements over time. With changes in the items used on the survey to reflect changes in screening patterns and potential interference from re-sampling, estimates cannot be directly compared across survey years. Further, NHIS relies on self-reported measures of screening use, which may have caused an overestimation of the true screen rates, particularly in some demographic groups.[251] Nevertheless, the survey provides one of the best estimates of the use of screening in the U.S.

The outcomes of the model strongly depend on the test performance assumptions for colonoscopy. There are currently no trial data available to validate the effectiveness of colonoscopy in our model.[86] Thus, we used adenoma miss rates from tandem colonoscopy studies to determine its efficacy,[163,252] and validated these estimates indirectly to outcomes of FOBT trials including colonoscopy follow-up of positive FOBT,[157,188,189,253,254] and the UK flexible sigmoidoscopy trial.[247] In our study screening alone was considered a sufficient explanation for the decrease in CRC mortality

between 1980 and 2000 and beyond (**Figure 5.2**). This may have overestimated the effects of screening; a previous microsimulation analysis suggested that treatment and risk factor developments also contributed to the decrease.[198] In a sensitivity analysis with a 50% reduced sensitivity for adenomas of ≤5 mm in diameter, the PAF for 2010 was lower than our base case estimate, but remained 52%. Under these assumptions, the age-adjusted relative risk for CRC mortality corresponding with colonoscopy screening was similar to the hazard ratio of 0.32 recently reported for a prospective cohort study with 22 years of follow-up.[166]

A limitation of using the PAF as a proxy for the potential returns of public health interventions is that it requires estimates of screening rates, an unproven constant estimate of the true magnitude of the benefit from screening and an approximation of the absolute disease risk in the population, all of which may change over time. Our estimates were based on currently available knowledge for each of these factors, but may not be applicable in future years, if more interventions to increase screening rates are implemented. We used PAF over other estimations such as the prevented fraction,[255,256] because the PAF metric can be used to provide policymakers estimates of potential future benefits of increased screening beyond current benefits of past exposure.

To conclude, a model-based approach estimated that more than half of the current CRC mortality risk in the U.S. is attributable to nonuse of screening. This underscores the need to increase screening uptake in the U.S. population. A model-based approach provided a higher estimate of screening benefit than the traditional, simpler approach to assess PAF. Valid estimation of the effects of screening requires the consideration of variable screening patterns over time, which may require more complex models than traditionally used to assess PAF.

APPENDIX 5

Supplementary Table 5.1 2010 National Health Interview Survey use and nonuse rates of colorectal cancer screening

Variable	Age strata (years)		
	50-64	65-74	75+
% Up-to-date (1-$P_{1;s}$) [a]	54.3%	67.9%	58.8%
% Not up-to-date ($P_{1;s}$)	45.7%	32.1%	41.2%

[a] The % of US adults age 50 years or older which had a colonoscopy in the last 10 years, a sigmoidoscopy in the last 5 years or a home fecal occult blood test in the last year.

Definition of PAF

We used the following symbols/acronyms:

s,t,i indicators for stratum, calendar year and screening test

PAF population attributable fraction

$pd_{s,t}$ the proportion of all risk events (deaths) occurring in stratum *s* and calendar year *t*

$P_{1;s,t}$ proportion exposed to nonuse of screening, by stratum and calendar year

$R_{T;s,t}$ outcome risk in the total population, by ,,

$R_{0;s,t}$ outcome risk in the absence of exposure, by ,,

$R_{1;s,t}$ outcome risk in the presence of exposure, by ,,

$RR_{x/y}$ risk ratio of situation x over y

PAF was defined by stratum and calendar year as:

$$\mathbf{PAF}_{s,t} = \frac{\mathbf{R}_{T;s,t}-\mathbf{R}_{0;s,t}}{\mathbf{R}_{T;s,t}} = \frac{\mathbf{RR}_{T/0;s,t}-1}{\mathbf{RR}_{T/0;s,t}} \quad (1)$$

Aggregate PAF per calendar year was obtained by weighing with the proportion of deaths occurring in each stratum:

$$PAF_t = \sum_s pd_{s,t} PAF_{s,t} \quad (2)$$

Levin's approach approximated (1) using the following formula:

$$PAF_{s,t} = \frac{P_{1;s,t}(RR_{1/0;s,t}-1)}{P_{1;s,t}(RR_{1/0;s,t}-1)+1} \quad (3)$$

$RR_{1/0;s,t}$ is the inverse of the risk (/rate) ratio corresponding with exposure to screening:

$$RR_{1/0;s,t} = \frac{1}{RR_{0/1;s,t}} \quad (4)$$

PART III

Potential Determinants of Screening Effectiveness

Chapter 6

Effectiveness of colonoscopy versus sensitive fecal occult blood screening for colorectal cancer; A microsimulation model based on National Colonoscopy Study data

Submitted:

RG Meester, I Lansdorp-Vogelaar, TR Church et al.

ABSTRACT

IMPORTANCE: Colonoscopy and fecal testing are recommended for colorectal cancer screening partly supported by modeling analyses with assumed 100% patient adherence. Differences in actual patient adherence may affect the long-term effectiveness of each test.

OBJECTIVE: Comparing the effectiveness of a program of ten-yearly colonoscopy versus annual sensitive guaiac-based fecal occult blood testing (FOBT) with observed longitudinal patient adherence.

DESIGN: Microsimulation model informed by observed National Colonoscopy Study data.

SETTING: Simulated randomized clinical trial from the United States

PARTICIPANTS: 3523 average-risk patients aged 40-69 years

EXPOSURES: We simulated a screening strategy of ten-yearly colonoscopy versus annual FOBT. Assumed adherence, FOBT positivity, and diagnostic colonoscopy adherence in FOBT-positive patients were based on observed National Colonoscopy Study data (≥4 FOBT rounds). For reference, we also simulated hypothetical scenarios of no screening and 100% screening adherence.

MAIN OUTCOMES: Estimated 15-year colorectal cancer incidence and mortality per 1000 patients with 95% probability intervals (95%PI) from multivariate probabilistic sensitivity analysis.

RESULTS: With no screening, the simulated incidence and mortality risks were 20.9 [95%PI, 15.8-26.9] and 6.9 [95%PI, 5.0-9.2] per 1000 patients, respectively. In the hypothetical case of 100% adherence, only colonoscopy was estimated to result in lower incidence (13.1 [95%PI, 9.7-17.0] for colonoscopy versus 20.9 [95%PI, 16.1-29.2] for FOBT), however, both tests lowered estimated mortality to a similar level (2.1 [95%PI, 1.6-2.9] versus 2.5 [95%PI, 1.9-3.5], respectively). Observed National Colonoscopy Study adherence levels were higher for colonoscopy (86%) than FOBT (80% completing at least one test), resulting in a larger loss in effectiveness for FOBT compared to screening with 100% adherence. Colonoscopy with observed patient adherence decreased estimated incidence to 14.2 [95%PI, 10.6-18.3] and mortality to 2.8 [95%PI, 2.1-3.8], while annual FOBT with observed patient adherence did not

influence estimated incidence [20.8, 95%PI, 15.9-27.7] and reduced mortality to 3.9 [95%PI, 3.0-5.5].

CONCLUSION: If patient adherence is as observed in NCS rather than assumed for current guideline recommendations, modeling suggests that colonoscopy may result in substantially greater reductions in colorectal cancer mortality than a program of annual FOBT.

INTRODUCTION

While over the past decades several independent randomized controlled trials have demonstrated that both invasive and fecal tests can be effective for colorectal cancer screening,[117,119,120,154-156,162,188] most currently recommended tests have not been directly evaluated in any trial.[86,122,257] Observational data suggest that screening colonoscopy exposure may have long-term preventive effects of 50-90%,[107] however, few comparable data exist on the effects of sensitive FOBT or fecal immunochemical testing (FIT).[105,258] United States Preventive Services Task Force screening recommendations to use colonoscopy or fecal testing for screening were partly informed by microsimulation modeling analyses with assumed 100% patient adherence for all tests.[259] Actual patient adherence rates for colonoscopy and fecal testing methods may differ and influence the long-term effectiveness of each test.[260] While fecal testing may have higher initial acceptance rates,[83-86] patients' willingness to comply with annual fecal colorectal cancer testing methods over longer periods of time is uncertain.[90-92,261]

In this study, we estimated the long-term colorectal cancer incidence and mortality effects for colonoscopy versus a program of annual FOBT using microsimulation modeling with adherence and outcome data from NCS.

METHODS

National Colonoscopy Study

NCS is a screening feasibility trial of colonoscopy versus a program of annual sensitive FOBT conducted in three clinical centers from geographically and demographically diverse areas in the United States. Participating centers include Group Health Cooperative (GHC) in Puget Sound, Washington, the University of Minnesota (MIN), Minneapolis, in Minnesota, and the Louisiana State University (LSU) in Shreveport, Louisiana. Participants for the study were recruited from current health plan members (GHC), health program participants (LSU, or mailing list members (UMN).

Participants for the study were recruited from current health plan members (GHC), health program participants (LSU), or mailing list members of the participating study centers between October 2004 and June 2008. Patients were aged 40-69 years at LSU, and 50-69 years in the two other centers. Earlier age for screening was instituted at LSU to allow for a pilot study in African Americans and whites of ages 40-49 years. Excluded were patients with a personal history of colorectal cancer, familial adenomatous polyposis, Lynch syndrome, or inflammatory bowel disease, as well as patients who had a prior colonoscopy or a flexible sigmoidoscopy within the last 5 years, or patients with serious comorbidities or an implanted defibrillator.

NCS recruited patients via a 2-step process. First, patients received an introductory letter with information about colorectal cancer, study intent and eligibility, and voluntary consent. Patients were also informed that participation in screening was free of charges. Next, those eligible and willing to participate were randomized to once-only screening colonoscopy, or a program of annual FOBT (see CONSORT flow chart). Randomization was conducted in a 1:1 fashion, with permuted blocks of varying sizes (2-6) for each study center developed by MSK. Those assigned to screening colonoscopy were contacted for scheduling by the clinical center; those assigned to annual FOBT were given the FOBT (Hemoccult SENSA) slides with instructions. Screening colonoscopy was offered no longer than one year. FOBT was offered up to 7 times. Patient navigators served both study arms comparably.

NCS data were collected by the individual study centers, but stored and analyzed centrally at MSK in New York. Outcomes used to inform this study include colonoscopy and FOBT adherence, FOBT positivity, diagnostic colonoscopy adherence, adenoma detection rates, advanced neoplastic lesions (adenomas ≥10mm in diameter, adenomas with high-grade dysplasia, or cancer), and cancers. Test adherence was defined as test completion within 1 year from each offering. FOBT positivity was assessed by two experienced laboratory technicians from the MSK Clinical Chemistry Laboratory. Colonoscopies were performed by board-certified endoscopists, and findings were histologically confirmed by an experienced pathologist at Boston University who was blinded to the exam indication and study arm.

Follow-up ended on the first of several events: cancer incidence, death, study close date (October 31, 2011), or other loss to follow-up. A full protocol of the study is enclosed as a supplementary file.

Microsimulation model

This study used the MISCAN-colon model (**Chapter 2**). In MISCAN-colon, the modeled effects of screening follow from a test's assumed ability to detect lesions within its reach or scope (**Supplementary Table 6.1**).[163,262] The simulated effects are concordant with randomized controlled trial data for screening with guaiac fecal occult blood tests[190] and sigmoidoscopy.[117] Estimated incidence and mortality effects of colonoscopy screening are consistent with the reported range in observational studies.[107]

Analysis

We simulated the NCS study population in terms of the age and sex distribution at enrolment. For patients in the colonoscopy arm, we simulated screening with an adherence rate equal to the observed overall colonoscopy completion rate in NCS. Colonoscopy was repeated after 10 years with similar assumed adherence. For patients in the FOBT arm, we simulated screening with long-term cumulative test adherence,

diagnostic colonoscopy adherence following a positive FOBT result, and potential crossover to screening colonoscopy derived from observed NCS data for up to 7 annual screening rounds. As the main determinant of FOBT positivity, assumed test specificity was varied to replicate observed positivity. Patients with detected adenomas received surveillance according to guidelines with assumed 80% adherence.[207]

For model validation purposes, simulated adenoma detection and cancer diagnosis during the study period were compared to the observed data. We then compared simulated long-term colorectal cancer outcomes between both study arms. For FOBT, we estimated outcomes both including and excluding colonoscopy crossover. For reference, we also estimated cancer incidence and mortality in hypothetical scenario of no and 100% adherence, to compare simulated screening benefits in the case of actual observed adherence to the maximum theoretical benefits in case of 100% adherence. Primary outcomes were simulated 15-year risks of colorectal cancer incidence and mortality.

Multivariate probabilistic analysis was conducted to derive 95% confidence intervals for all base-case model outcomes. We varied 16 key parameters along uniform, beta, or lognormal distributions in 1000 simulation runs of 10 million persons (**Supplementary Table 6.1**).[263]

Sensitivity analysis

In sensitivity analyses, we re-assessed outcome differences for colonoscopy and FOBT successively excluding patients who did not complete the first offered test and assuming 100% adherence with follow-up colonoscopy of positive FOBTs; separately evaluating outcomes for each participating study center; for once-only colonoscopy screening; assuming 50% lower colonoscopy sensitivity for diminutive adenomas (in accordance with previous estimates);[263] assuming lower FOBT specificity;[259] applying estimated FIT performance characteristics for FOBT;[259] and, excluding patients aged 40-49 years.

Study oversight

The institutional review board of each study center participating in NCS approved the study, and ascertained informed consent for included patients. The study was coordinated by MSK. MISCAN-colon was developed and employed for this study at the Erasmus MC Department of Public Health, Rotterdam, Netherlands. This study is funded by the United States National Cancer institute.

RESULTS

A total of 3523 patients were enrolled in NCS, of whom 1761 were assigned to screening colonoscopy and 1762 were assigned to annual FOBT screening (**Figure 6.1**). Patients

in both study arms were comparable in terms of age (mean 55, SD 5.5), gender (50% versus 48% male), and race/ethnicity (81% Caucasian) (**Supplementary Table 6.2**).

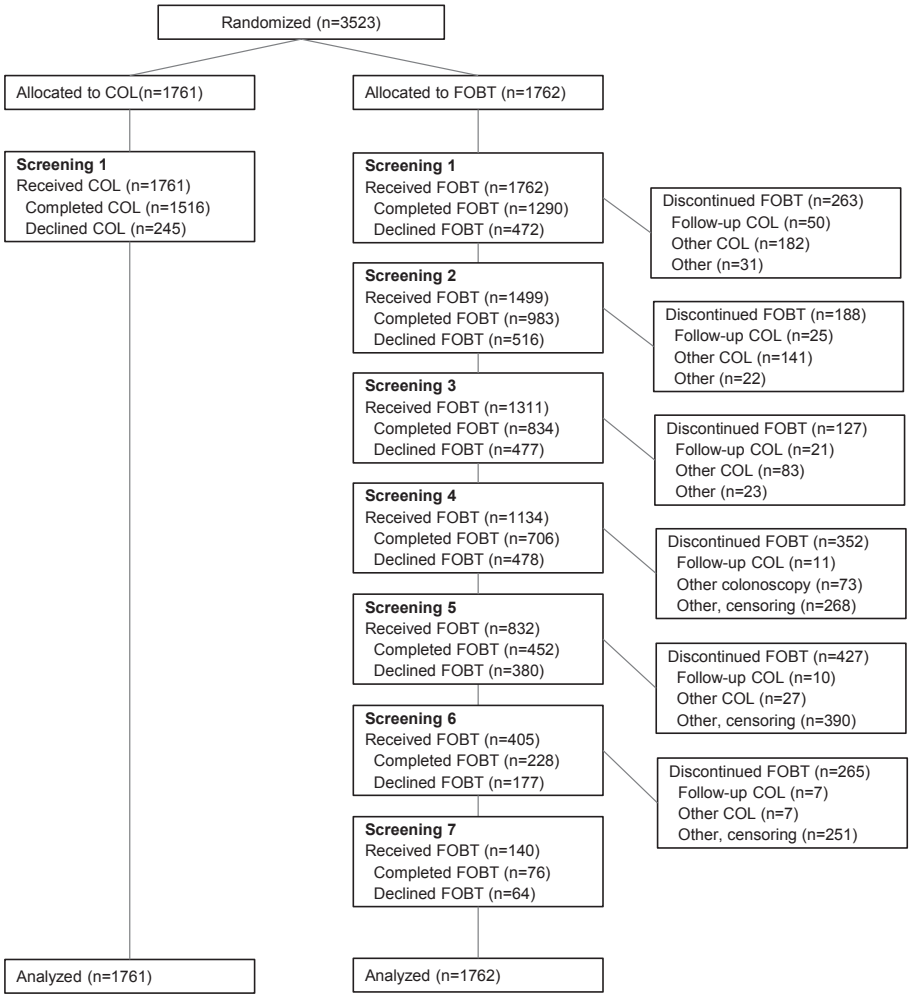


Figure 6.1. National Colonoscopy Study flow diagram (CONSORT)

Of the 1761 patients randomly assigned to colonoscopy screening, 1516 (86%) completed the examination (**Figure 6.1**, **Supplementary Table 6.3**). Of the 1762 patients assigned to FOBT screening, 1290 (73%) completed the first test; 1184 patients were offered ≥4 FOBT, and of these 948 (80%) patients completed at least one test, 840 (71%) completed at least two, and 585 (50%) completed all four tests; a subset of 140 early study participants were offered 7 tests, and of these 119 (85%) completed at least one test, and 61 (46%) completed all 7 tests. Positive tests varied

from 51 (4.0% of completed FOBTs) in round 1, to 11 (2.0%) in round 4, to 3 (4.0%) in round 7. Of all 139 patients with a positive test throughout the study, 127 (91%) completed a diagnostic colonoscopy. A total of 513 (29%) patients in the FOBT arm received colonoscopies for reasons other than a positive FOBT result.

Model calibration and validation

The MISCAN-colon model replicated observed colonoscopy adherence and cumulative FOBT adherence, as well as observed FOBT positivity for 7 rounds (**Figure 6.2**). Simulated short-term adenoma findings were higher than observed NCS findings (FOBT non-significantly). Simulated advanced adenoma detection rates and cancer diagnosis rates were within or on 95% probability bounds around the observed (**Supplementary Table 6.4**).

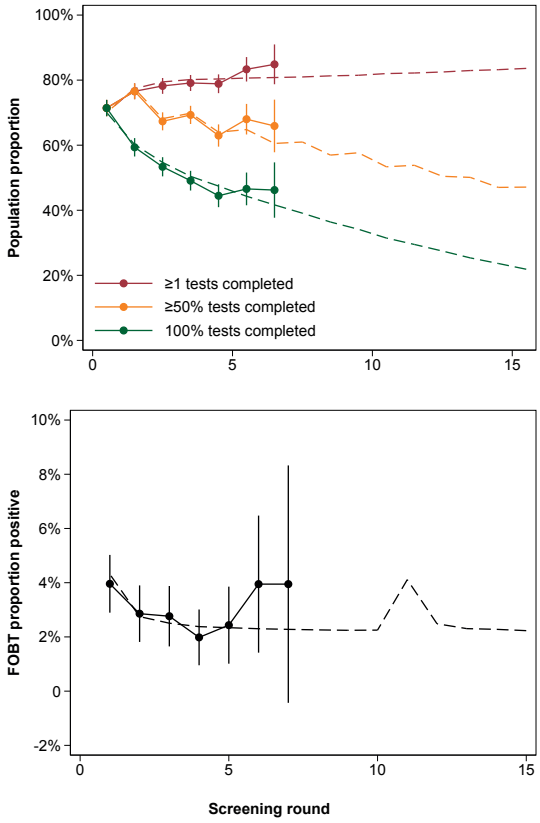


Figure 6.2a-b. Observed (solid) versus simulated (dashed lines) FOBT adherence (a) and positivity (b) [a]

[a] Analysis here excludes patients who crossed over to colonoscopy. Round 6-7 data represent one study center.

Cancer incidence and mortality

Without any screening, the model estimated that the 15-year colorectal cancer incidence risk for the NCS study population was 20.9 [95%PI, 15.8-26.9] per 1000. The estimated cancer mortality risk without screening was 6.9 [95%PI, 5.0-9.2] per 1000.

Colonoscopy screening with observed NCS adherence rates was estimated to decrease colorectal cancer risk to 14.2 [95%PI, 10.6-18.3] per 1000 (-32% compared to no screening), and mortality risk to 2.8 [95%PI, 2.1-3.8] per 1000 (-59%) (**Figure 6.3a**). FOBT screening with observed NCS adherence rates without crossover colonoscopy was estimated in the model to result in similar cancer risk of 20.8 [95%PI, 15.8-28.1] per 1000 (-1% compared to no screening), however, it decreased estimated cancer mortality to 3.9 [95%PI, 2.9-5.4] per 1000 (-43%) (**Figure 6.3b**). Including crossover colonoscopy, screening in the FOBT arm reduced simulated incidence and mortality to 19.2 [95%PI, 14.1-23.4] per 1000 (-8%) and mortality 3.5 [95%PI, 2.6-4.5] per 1000 (-49%), respectively. The estimated relative risk of colorectal cancer mortality in those screened with colonoscopy versus those screened with FOBT was 0.72 [95%PI, 0.65-0.77].

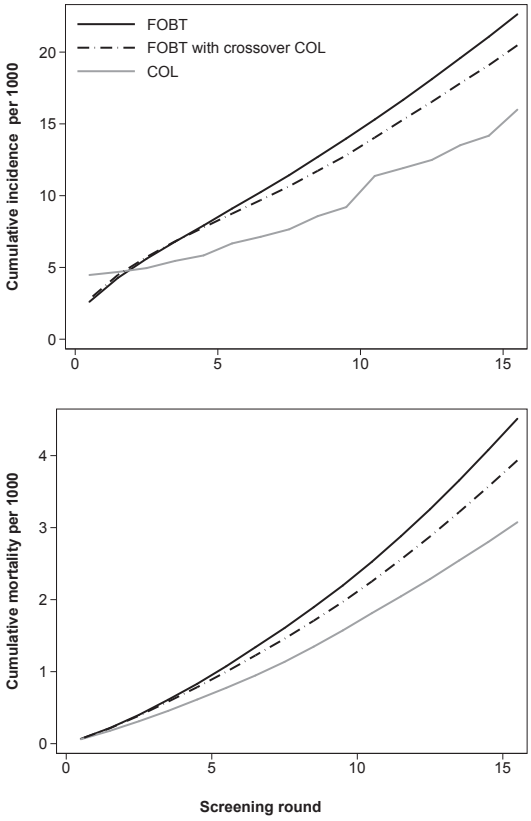


Figure 6.3a-b. Simulated colorectal cancer incidence and mortality in two screening strategies
Abbreviations: COL = colonoscopy.

Loss in effectiveness from incomplete adherence

With 100% assumed patient adherence, colonoscopy screening was estimated to result in 15-year incidence and mortality risks of 13.1 [95%PI, 9.7-17.0] per 1000 (-38% compared to no screening) and 2.1 [95%PI, 1.6-2.9] per 1000 (-69%), respectively, while FOBT screening resulted in risks of 20.9 [95%PI, 15.8-29.4] per 1000 (-0%) and 2.5 [95%PI, 1.8-3.4] per 1000 (-64%). Compared to screening with full patient adherence, the actual use of colonoscopy screening was associated with a relative loss of 14% in screening effectiveness to reduce 15-year cancer mortality risks (**Figure 6.4a**). FOBT without crossover colonoscopy was associated with a relative loss of 33% (22%, including colonoscopy crossover) (**Figure 6.4b**).

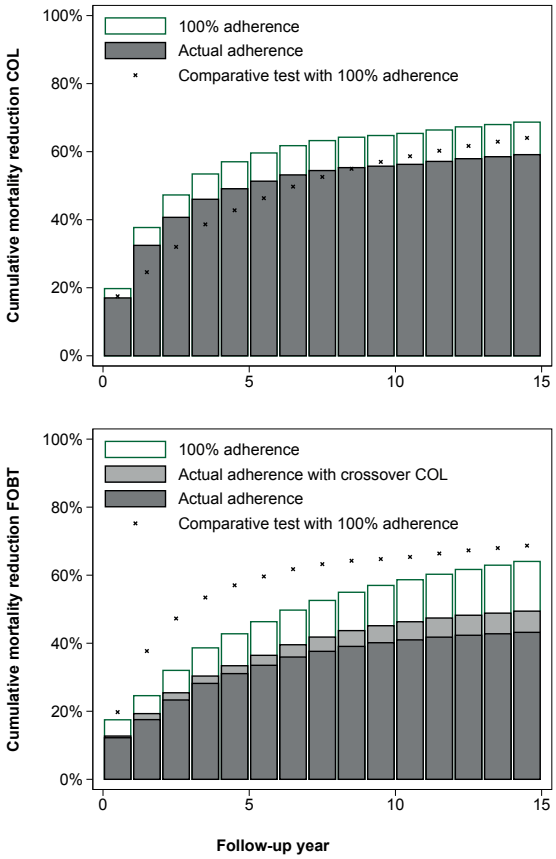


Figure 6.4a-b. Loss in estimated screening benefit with actual versus 100% patient adherence [a]
Abbreviations: COL = colonoscopy.
[a] The comparative test is FOBT in Figure 6.4a and colonoscopy in Figure 6.4b.

Sensitivity analysis results

Estimated relative mortality benefits compared to no screening increased to 69% for colonoscopy versus 56% for FOBT when including only attenders for colonoscopy and initial FOBT (absolute risks 2.1 versus 3.0 per 1000) (**Table 6.1**). Estimated mortality benefits varied less across participating study centers (**Supplementary Table 6.5**), with once-only colonoscopy screening, for higher ages, or with alternative diagnostic test performance assumptions.

Table 6.1. Simulated colorectal cancer mortality risks across sensitivity analyses

	Screening modality				
	None	Colonoscopy		FOBT	
Analysis		Risk	% Diff	Risk	% Diff
1. Base-case	6.9	2.8	-59%	3.9	-43%
2. First round attenders only [a]		2.1	-69%	3.0	-56%
3. By center					
a. GHC		2.9	-58%	4.0	-42%
b. LSU		3.1	-54%	4.3	-38%
c. MINN		2.6	-62%	3.4	-50%
4. Once-only colonoscopy		3.0	-56%	3.9	-43%
5. Lower COL sensitivity [b]		3.1	-54%	4.1	-40%
6. Lower FOBT specificity [c]		2.8	-59%	3.7	-47%
7. FIT characteristics [d]		2.8	-59%	3.5	-48%
8. Age ≥50y	7.4	3.0	-59%	4.1	-45%

Abbreviations: GHC = Group Health Cooperative; LSU = Louisiana State University; MINN = University of Minnesota; FIT = fecal immunochemical test.

[a] Only patients who completed the first offered test were included in the analysis. Assumed diagnostic colonoscopy adherence was 100%.

[b] Assumed colonoscopy sensitivity for small adenomas up to 5 mm in diameter was 50% lower than the base-case.

[c] Assumed FOBT specificity is similar to US Preventive Services Task Force analysis.[259]

[d] Assuming similar test use over time, we applied FIT characteristics instead of FOBT performance characteristics.[259]

DISCUSSION

We used NCS adherence and outcome data with microsimulation modeling to estimate the long-term colorectal cancer effects for colonoscopy versus a program of annual sensitive FOBT. The observed proportion of patients completing screening colonoscopy was 86% versus only 50% for FOBT who completed all tests after 4 rounds and 80% who completed ≥1. Although these adherence levels were associated in the model with substantial estimated mortality reductions for both tests, the

disparity in adherence translated in a greater mortality reduction for colonoscopy compared to FOBT, and a smaller loss in mortality benefit compared to screening with 100% patient adherence of 14% for colonoscopy versus 33% for FOBT.

In the NCS, there was frequent colonoscopy use in the FOBT arm. Of the patients assigned to annual FOBT, 29% received a colonoscopy outside the study. Indications for the exams were unknown, but likely, most of them were opportunistic screening exams. Considering opportunistic colonoscopy use in the FOBT arm, the overall proportion of patients who completed any screening test was approximately similar to that in the colonoscopy arm. When we included colonoscopy crossover in our analyses, the estimated outcomes differences between colonoscopy and FOBT screening were smaller (50% vs 59% mortality reduction).

The estimated effectiveness differences between colonoscopy and FOBT were robust for a number of factors evaluated in sensitivity analyses. First, we adjusted the result comparison for disparities in nonuse of screening. In colonoscopy screening, 14% of all patients had not completed the exam, while in annual FOBT, of the patients offered at least 4 tests, approximately 20% had not completed any of the offered tests. Adherence with diagnostic follow-up in case of positive FOBT results was also less than 100%. In a comparison of patients who completed colonoscopy versus those who completed at least one FOBT and potential diagnostic colonoscopy, the effectiveness differences were not influenced substantively (**Table 6.2**). This suggests that persistent high FOBT adherence is required to observe more similar benefits for FOBT and colonoscopy.

Further, despite differences in adherence and estimated outcomes across participating study centers, colonoscopy resulted in substantially greater estimated mortality reductions for all centers (**Supplementary Table 6.5**). Once-only colonoscopy screening had nearly similar estimated outcomes as colonoscopy screening every 10 years over a 15 year risk period. Results were also robust for alternative test performance assumptions. Previous studies have suggested that our base-case assumptions for colonoscopy sensitivity for small adenomas may be overly optimistic,[263] potentially overstating the effects of colonoscopy screening. Similarly, assumed FOBT specificity was higher in this study than in previous studies to accommodate for the relatively low observed positivity rate in the trial,[259] which may have underestimated the use of colonoscopy and chance findings in FOBT screening, and therefore the overall benefits of FOBT. Further, many settings use FIT instead of FOBT, which may have more favorable performance characteristics.[262] However, we found only modest effects on relative mortality reductions for each of the assumptions (maximum 5% point). Finally, exclusion of patients under 50 years of age (n=370), for whom regular screening generally is not recommended in the United States,[264] also did not affect the study conclusions.

Compared to previous randomized clinical trials investigating comparative test adherence, observed colonoscopy adherence in NCS was relatively high. Most studies directly comparing colonoscopy and FOBT or FIT in a single round reported higher adherence rates for stool-based tests.[83-86] The participation rates in our study reflect those of patients providing consent to participate in screening with either colonoscopy or FOBT. Further, colonoscopy costs were completely covered within the context of this study, which may have given patients an incentive to get a colonoscopy within the study setting. This suggests that our results for the comparison of colonoscopy versus FOBT are applicable primarily to patients who are willing to undergo colonoscopy screening. Although this may represent a limited subgroup of the total population for some settings, in the United States, screening colonoscopy is widely used for screening.[89]

In contrast to colonoscopy adherence rates, FOBT participation rates in our study were similar to observed adherence rates for previous low-sensitivity guaiac fecal occult blood testing trials.[188,265] Interestingly, they were also comparable to cumulative FIT adherence rates from recent population-based studies (for ≤4 rounds) with systems in place to track and remind non-adherent patients.[90-92,261] Many national programs are known to have much lower population adherence rates.[5] The consistency with observational studies suggests that our results for FOBT may be generalizable to other settings with patient tracking and reminding. In settings without such services, outcomes of screening in general may be less favorable.

A recently published comparative modeling analysis for the U.S. Preventive Services Task Force (USPSTF) found more similar effects for colonoscopy and FOBT.[259] Our analysis differs from that study in two important ways. First, the analysis for the USPSTF assumed 100% patient adherence for both tests. In general, modeling studies assume either full, or a fixed lower level of patient adherence with screening.[111,123,211,266] To our knowledge, the present study is the first to closely replicate long-term observed test adherence patterns for FOBT in order to assess the associated benefits. As we showed, the observed adherence disparity for colonoscopy and FOBT had profound implications for the screening benefits.

Another difference with the analyses used to inform the USPSTF is our use of longitudinal FOBT positivity data to derive test performance assumptions. The observed test positivity rate in the initial round was 4%, which is substantially lower than the 7.6% assumed false positivity rate elsewhere.[259] Positivity rates decreased further after the first round to approximately 2.5%, similar to a another report.[267] The decline in test positivity could not be explained entirely in the model by higher first round prevalence of cancer and adenoma cases, and thus, we assumed that the rate of false positivity decreased over time similar to another recent study.[268] This reduced the number of colonoscopy examinations for asymptomatic blood loss. Although the

sensitivity analysis indicated that this has only modest health outcome effects, it may influence the efficiency of the test.

Similar to the recent USPSTF analysis,[259] we modeled the effects of colonoscopy and sensitive FOBT screening by combining observational data on their diagnostic performance,[163] with randomized controlled trial data on flexible sigmoidoscopy,[117] and low-sensitivity guaiac-based FOBT screening.[154,156,188] Our simulated short-term advanced adenoma and cancer detection rates were consistent with observed NCS data (**Supplementary Table 6.4**). We overestimated overall adenoma detection rates for both colonoscopy and FOBT (cancer detection rates non-significantly). This may suggest that the NCS study population had lower-than-average risk, as observed before for the Minnesota Colorectal Cancer Control Study.[156] In contrast to the Minnesota study, we found no preventive effect for annual guaiac FOBT screening, despite a higher sensitivity of Hemoccult Sensa for adenomas compared to the Hemoccult II test. We assumed in our study that patients received no screening prior to participating in the study, which increased detection of cancer cases in initial screening years (i.e. prevalence screening rounds). With longer simulated follow-up (lifetime), we did find a substantial incidence reduction in line with Minnesota trial results (*results not shown*).[259]

NCS is the first study to assess FOBT adherence and outcomes during more than 4 subsequent rounds. The model was able to accurately replicate data for up to 5 rounds. A limitation of the study is that there was a relatively high rate of loss to follow-up for rounds 5-7. Observed data for round 6 and 7 represent a single study center out of three total participating study centers, and may be less representative therefore of the general population. To minimize the potential bias from loss to follow-up, our microsimulation model gave the most weight to earlier years in fitting and projecting FOBT adherence and positivity (**Figure 6.2**).

To conclude, with observed patient adherence data from NCS rather than assumptions used in current guideline recommendations, colonoscopy results in a substantially greater reduction in colorectal cancer incidence and mortality than a program of annual FOBT. These results imply that if patients are willing to undergo colonoscopy, this may result in superior outcomes to annual FOBT due to likely suboptimal long-term adherence for the latter test. In offering stool-based tests to average-risk patients age 50 and older, guidelines need to emphasize the importance and effect of high patient adherence.

APPENDIX 6

Supplementary Table 6.1. Key modeling assumptions

	Base-case value	PSA value	Refs
Demography			
All-cause mortality	U.S. lifetables		[176]
Natural History			[g]
Adenoma onset	Nonhomogeneous Poisson process		[181,182,185]
	Exponential(λμ) time to event		
	λ~Gamma(1;2) risk dispersion factor	Unif(-20%+20%)	
	μ=260-18 for age 25-80y	Unif(-10%+10%)	
Adenoma progression			[117]
State transitions	0-89% progressive for age 0-100y	Unif(-10%+10%)	
	30% size 6-9mm progress to cancer		
	70% size 6-9mm first become 10+mm		
State durations, y (total)	Exponential(λ=130)	λ~Unif(-10%+10%)	
Preclinical cancer progression			[157,188,189]
Stage transitions [a]	0-31% stI become clinical for age 0-100y	Unif(-10%+10%)	
	18-58% stII become clinical for age 0-100y	Unif(-10%+10%)	
	58-49% stIII become clinical for age 0-100y	Unif(-10%+10%)	
Stage durations, y (average)	Exponential(2.5)	Unif(-10%+10%)	
Colorectal cancer incidence (without exposure to screening)	See Figure 2.3-4		SEER 1975-79 [175]
5y Colorectal cancer survival [b]	58-71% stI, depending on location		SEER 2000-10 [175]
	58-62% stII, depending on location		
	33% stIII		
	6% stage IV		
Colonoscopy performance			
Sensitivity, % [b]			
adenomas 0-5mm	75	Beta;SE:3.5	
adenomas 6-9mm	85	Beta;SE:3.5	[163]
adenomas ≥10mm	95	Beta;SE:2.5	[163]
malignant neoplasia		Beta;SE:2.5	
Specificity, % [c]	100	Beta;SE:5	[269,270]
Complete colonoscopy examination, % [d]	95	Beta;SE:2.5	
Complication rates, %			
with polypectomy	Age-dependent (50-100 years):		[271,272]
Serious GI complications	0.2-2.9	LogN;SE:10%	

Supplementary Table 6.1. Key modeling assumptions

	Base-case value	PSA value	Refs
Fatal complications	0.0033	LogN;SE:50%	
Other GI complications	0.2-2.6	LogN;SE:10%	
CV complications	0.1-2.5	LogN;SE:10%	
without polypectomy [e]	-		
FOBT performance			[259]
Sensitivity, %			
adenomas 0-5mm	0		
adenomas 6-9mm	4.3	Beta;SE:2.5	
adenomas ≥10mm	14.7	Beta;SE:3.5	
malignant neoplasia [f]	56.8 / 85.9	Beta;SE:3.5	
Specificity, % [c]	97.1	Beta;SE:3.5	

Abbreviations: PSA= probabilistic sensitivity analysis; CDC = U.S. Centers for Disease Control and Prevention; Poisson = Poisson distribution; Unif = uniform distribution; Exp = exponential distribution; SEER = Surveillance Epidemiology and End Results program; Beta = beta distribution; SE = standard error; GI = gastrointestinal; LogN = lognormal distribution.

[a] In multiway probabilistic sensitivity analyses the model parameters were varied randomly according to Uniform, Beta or Lognormal distributions. To limit the degrees of freedom, several parameters were assumed to be perfectly correlated: adenoma onset related parameters, adenoma progression-related variables, cancer progression-related variables, sensitivity for small adenomas with sensitivity for medium adenomas, sensitivity for large adenomas with sensitivity for cancer, all complication types. [b] Sensitivity was defined as the probability of detecting an adenoma that was present at the time of exam. Based on baseline-detection rates in our data, sensitivity for colorectal cancers was assumed to be unrelated to ADR.

[c] The occurrence of false positive FOBT results was assumed non-random for some patients. We assumed perfect specificity for colonoscopy including pathological examination of detected lesions.

[d] Colonoscopy was considered complete if the cecum was reached. In the 2% incomplete examinations, the endpoint was assumed to be distributed uniformly over the colon/rectum.

[e] Colonoscopy without polypectomy was not associated with a higher risk of complications. The risk of complications for polypectomy increased exponentially with age. Complications include serious GI events such perforation and gastrointestinal bleeding requiring blood transfusions; other GI events such as paralytic ileus, nausea, vomiting and dehydration, abdominal pain; and cardiovascular events such as myocardial infarction or angina, arrhythmias, congestive heart failure, cardiac or respiratory arrest, or syncope, hypotension, or shock. The fatal perforation rate was derived from estimates of the incidence of perforation and case-fatality for perforation.[271 272]

[f] We assumed that fecal occult blood testing is more sensitive in cancers towards the end of the occult invasive period (close, time-wise, to becoming symptomatic): for preclinical cancers which will become symptomatic within the same stage, assumed test sensitivity was higher. This assumption showed good concordance with guaiac fecal occult blood test trial results.[190]

[g] More details regarding the calibrated natural history parameters and other model elements are provided in **Chapter 2**.

Supplementary Table 6.2. Baseline patient characteristics

	Colonoscopy	FOBT	Total
Total patients, n	1761	1762	3523
Demographics			
Average age, y (SD)	55 (5.50)	55 (5.5)	55(5.54)
Men, n (%)	845 (48%)	881 (50%)	1726 (49%)
White, n (%)	1426 (81%)	1428 (81%)	2854 (81%)
College graduate, n (%)	704 (40%)	634 (36%)	1304 (37%)
Risk factors			
Obese, (Body Mass Index $\geq$ 30) n (%)	511 (29%)	581 (33%)	1092 (28%)
Regular multivitamin use, n (%)	881 (50%)	793 (45%)	1674 (48%)
Aspirin use, n (%)	546 (31%)	511 (29%)	1057 (30%)
Hormone use (Women), n (%)	652 (37%)	652 (37%)	1304 (37%)
Current Smoker, n (%)	211 (12%)	211 (12%)	422 (12%)
Family history (First Degree Relatives with CRC), n (%)	158 (9%)	142 (8%)	300 (8%)

Supplementary Table 6.3. National Colonoscopy Study adherence and outcomes

Outcome	Colonoscopy	FOBT						
		Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7
Number invited / mailed	1761	1762	1499	1311	1184	832	405	140
Adherence (marginal)	1516 (86%)	1290 (73%)	983 (66%)	834 (64%)	706 (60%)	452 (54%)	228 (56%)	76 (54%)
Adherence (cumulative) [a]	n.a.	1290 (73%)	914 (61%)	716 (55%)	585 (50%)	362 (45%)	180 (46%)	61 (46%)
Cross-over (marginal) [b]	n.a.	182 (10%)	141 (9%)	83 (6%)	73 (6%)	27 (3%)	7 (2%)	-
Cross-over (cumulative) [c]	n.a.	182 (10%)	323 (18)	406 (23%)	479 (27%)	506 (29%)	513 (29%)	513 (29%)
Positive tests	n.a.	51 (4%)	28 (2.8%)	23 (2.8%)	14 (2.0%)	11 (2.4%)	9 (4.0%)	3 (4.0)
Adherence to Dx COL	n.a.	50 (98%)	25 (89%)	21 (91%)	11 (79%)	10 (90%)	7 (78%)	3 (100%)
DR NAA [d]	344 (20%)	189 (11%)						
DR AA [d]	83 (5%)	73 (4%)						
Cancers [d]	3 (0.2%)	6 (0.5%)						

Dx COL = Diagnostic colonoscopy after a positive FOBT; DR = detection rate; NAA = non-advanced adenoma; AA = advanced adenoma.

[a] Proportion of patients having returning all mailed FOBT from to the total number invited up to that round.

[b] Proportion of patients assigned to FOBT who had a screening colonoscopy outside the study without having a preceding positive FOBT result.

[c] Cumulative proportion of patients crossing over relative to the total number of enrolled patients.

[d] Only screen-detected findings included, for FOBT across all rounds.

Supplementary Table 6.4. Simulated versus observed adenoma detection and cancer diagnosis in the National Colonoscopy Study [a]

Finding	Colonoscopy			FOBT		
	Observed		Simulated	Observed		Simulated
	Mean	95% CI		Mean	95% CI	
Adenomas, %	19.5	17.7-21.4	23.3	3.8	2.9-4.7	4.8
Advanced adenomas, %	4.7	3.7-5.7	4.6	2.3	1.6-3	2.6
Cancer, %	0.2	0-0.4	0.4	0.5	0.2-0.8	0.8

[a] This comparison reflects screen-detected findings.

Supplementary Table 6.5. Adherence by National Colonoscopy Study center

Study center	Colonoscopy	FOBT						
		Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7
GHC	198/233 (85%)	180/235 (77%)	151/214 (71%)	146/200 (73%)	136/180 (75%)	52/83 (63%)	34/52 (65%)	8/15 (53%)
LSU	399/504 (79%)	310/503 (62%)	229/438 (52%)	196/410 (48%)	154/381 (40%)	117/295 (40%)	54/116 (47%)	11/27 (41%)
UMN	919/1024 (90%)	800/1024 (78%)	603/847 (71%)	492/701 (70%)	416/623 (67%)	283/454 (62%)	140/237 (59%)	57/98 (58%)
Total	1516/1761 (86%)	1290/1762 (73%)	983/1499 (66%)	834/1311 (64%)	706/1184 (60%)	452/832 (54%)	228/405 (56%)	76/140 (54%)

Abbreviations: GHC = Group Health Cooperative; LSU = Louisiana State University; UMN = University of Minnesota

Chapter 7

Variation in adenoma detection rate and the lifetime benefits and cost of colorectal cancer screening; A microsimulation model

Article published (AltMetric = 135):

RG Meester, CA Doubeni, I Lansdorp-Vogelaar et al.
JAMA. 2015 Jun 16;313(23):2349-58

ABSTRACT

IMPORTANCE: Colonoscopy is the most commonly used colorectal cancer screening test in the United States. Its quality, as measured by adenoma detection rates, varies widely between physicians with unknown consequences for the cost and benefits of screening programs.

OBJECTIVE: To estimate the lifetime benefits, complications and costs of a colonoscopy screening program at different levels of adenoma detection.

DESIGN, SETTING and PARTICIPANTS: This study used microsimulation modeling with data from a community-based healthcare system on adenoma detection rate variation and cancer risk among 136 physicians and 57,588 patients for 1998-2010.

EXPOSURE: Using modeling, no screening was compared to screening initiation with colonoscopy according to adenoma detection rate quintiles (averages 15.3, 21.3, 25.6, 30.9, and 38.7%) at ages 50, 60 and 70 with appropriate surveillance of adenoma patients.

MAIN OUTCOMES: Estimated lifetime colorectal cancer incidence, mortality, number of colonoscopies, complications and costs per 1,000 patients, all discounted at 3% per year and including 95% confidence intervals from multiway probabilistic sensitivity analysis (95%CI).

RESULTS: In simulation modeling, among unscreened patients, the lifetime risks of colorectal cancer incidence and mortality were 34.2 (95%CI:25.9-43.6) and 13.4 (95%CI:10.0-17.6) per 1,000, respectively. Among screened patients, simulated lifetime incidence decreased with lower to higher adenoma detection rates (quintile 1 versus 5: 26.6, 95%CI:20.0-34.3 versus 12.5, 95%CI:9.3-16.5) as did mortality (5.7, 95%CI:4.2-7.7 versus 2.3, 95%CI:1.7-3.1). Compared to quintile 1, simulated lifetime incidence and mortality were on average 11.4% (95%CI:10.3-11.9) and 12.8% (95%CI:11.1-13.7) lower, respectively, for every 5 percentage-point higher adenoma detection rate. Total colonoscopies and associated complications were higher from quintile 1 (2,777, 95%CI:2,626-2,943 and 6.0, 95%CI:4.0-8.5) to subsequent quintiles (quintile 5: 3,376, 95%CI:3,081-3,681 and 8.9, 95%CI:6.1-12.0). Estimated net screening costs were, however, lower from quintile 1 (US $2.1 million, 95%CI:1.8-2.4) to quintile 5 (US$1.8 million, 95%CI:1.3-2.3) due to averted cancer treatment costs. Results were stable across sensitivity analyses.

CONCLUSIONS-RELEVANCE: Using microsimulation modeling, we found that higher adenoma detection was associated with lower lifetime colorectal cancer incidence and mortality without higher overall costs. Future research is needed to assess if increasing adenoma detection would be associated with improved patient outcomes.

INTRODUCTION

Screening colonoscopy reduces colorectal cancer mortality risk through detection and treatment of precursor adenomatous or early cancerous lesions,[165-167] but its effectiveness depends upon exam quality.[163,273,274] A currently recommended colonoscopy quality indicator, the adenoma detection rate (ADR), has been found to vary at least 3-fold across physicians.[124,128,275] A recent large United States study found that this variation is associated with patient outcomes: compared to patients of physicians with the highest ADRs, patients of physicians with the lowest ADRs had a nearly 50% higher risk of colorectal cancer and a 60% higher risk of fatal disease during up to 10 years of follow-up after colonoscopy.[124] This suggests that higher adenoma detection is associated with both better disease detection and disease management. However, little is known about the consequences of different levels of ADR for the lifetime benefits, risks and cost in a program using colonoscopy as the initial and primary screening test in an average-risk population. Higher ADRs may accrue mostly from increased detection of small low-risk polyps, resulting in an increased number of subsequent surveillance colonoscopies and complications for polyps that may never cause fatal disease. Thus, any benefits of higher ADR may be outweighed by the corresponding harms.[141]

In the present study, we evaluated various outcomes for a colonoscopy-based colorectal cancer screening strategy according to different adenoma detection rate levels, including lifetime colorectal cancer incidence and mortality, the number of colonoscopies and related complications, and screening and treatment costs.

METHODS

We used microsimulation modeling of screening in a United States population cohort with community-based data on ADR variation and cancer risk. This study was approved by the Kaiser Permanente Northern California (KPNC) institutional review board, and conducted as part of the United States National Cancer Institute (NCI)-funded consortium Population-Based Research Optimizing Screening through Personalized Regimens (PROSPR), which aims to conduct multi-site, coordinated, trans-disciplinary research to evaluate and improve cancer screening.

KPNC data

Physician-level (ADR) and patient-level (age, sex, race/ethnicity, cancer diagnosis) data were from KPNC, an integrated healthcare delivery system.[124] The data for this study were confined to screening colonoscopies performed by 136 gastroenterolo-

gists between January 1, 1998 and December 31, 2010. Outcomes were ascertained in the 6-month to 10-year period after initial colonoscopy through December 31, 2010. The screening indication excluded patients who had prior: adenomas or colorectal cancer; inflammatory bowel disease within 10 years; colonoscopy within 10 years, sigmoidoscopy within 5 years; positive fecal hemoglobin test within 1 year; or abdominal symptoms within 6 months. ADRs, the proportion of a physician's screening colonoscopies that detect ≥1 histologically confirmed adenomas, ranged from 7.3% to 52.5%; the averages (and ranges) for ADR quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

Natural history of colorectal cancer

The MISCAN-colon model assumes that colorectal cancer develops progressively from small (≤5mm) through medium (6-9 mm) or large adenomas (≥10mm) (**Chapter 2**). An early stage tumor may progress to an advanced-stage tumor without symptoms, or become symptomatic during any stage and be clinically diagnosed. Some patients die of the disease and lose life-years, while others die from competing causes before or after developing cancer. Serrated adenomas are not modeled distinct from conventional adenomas.[56]

Performance characteristics of colonoscopy

The modeled effectiveness of colonoscopy screening depends on assumptions regarding its completeness and sensitivity for adenomatous lesions (**Table 7.1**). For this study, we used observed data from KPNC to derive sensitivities for colonoscopy at the five ADR quintiles, while assuming no underlying differences in adenoma prevalence.[276]

Table 7.1 Key modeling assumptions.

Input parameter	Base-case assumption	PSA assumption[a]	References
Demography			
All-cause mortality	U.S. Lifetables		CDC 2010[h]
Natural history			
Adenoma onset	Age-dependent (non-homogeneous Poisson)	Unif(-20%+20%)	[h]
Adenoma progression			
State transitions	Age-dependent	Unif(-10%+10%)	[h]
State duration, years (total)	Exp(λ=130)	λ~Unif(-10%+10%)	[h]
Cancer progression (preclinical)			
Stage transitions	Age-dependent	Unif(-10%+10%)	[h]
Stage durations, years	Exp (λ=2.5)	λ~Unif(-10%+10%)	[h]
Colorectal cancer incidence (without exposure to screening)	Age-/Stage-/Location-dependent		SEER 1975-79[h]
Colorectal cancer survival	Age-/Stage-/Location-dependent		SEER 2000-10[h]
Colonoscopy quality			
Sensitivity, % [b]	ADR quintile-dependent:		
adenomas 0-5mm	14.7-29.6-41.0-66.2-98	Beta;SE:3.5	[h]
adenomas 6-9mm	39.6-65.8-85.0-94.3-98	Beta;SE:3.5	[h,163]
adenomas ≥10mm	88.0-92.2-95.0-96.8-98	Beta;SE:2.5	[h,163]
malignant neoplasia	98	Beta;SE:2.5	[h]
Specificity, % [c]	85	Beta;SE:5	[269,270]
Complete colonoscopy examination, % [d]	98	Beta;SE:2.5	[286,287]
Complication rates, %			
with polypectomy	Age-dependent (50-100 years):		[271,272]
Serious GI complications	0.2-2.9	LogN;SE:10%	
Fatal complications	0.0033	LogN;SE:50%	
Other GI complications	0.2-2.6	LogN;SE:10%	
Cardiovascular complications	0.1-2.5	LogN;SE:10%	
without polypectomy [e]	-		
Costs, US $ [f]			
Colonoscopy			CMS 2007[152]
without polypectomy	899	LogN;SE:5%	
with polypectomy	1,140-1,270 for ADR q1-5	LogN;SE:5%	
Complication	6,129		CMS 2007[277]
Per life-year with cancer care [g]	Stage-dependent (I-IV):		CMS 2007[278]
Initial year, stage I-IV	37,185-78,876	LogN;SE:1.1-1.9%	
Ongoing, stage I-IV	3,092-12,350	LogN;SE:4.4-5.7%	
Terminal year, stage I-IV	64,693-89,600	LogN;SE:1.2-2.2%	
Terminal year, stage I-IV	19,427-50,552	LogN;SE:8.4-10%	

← **Table 7.1 Legend** Key modeling assumptions.
Abbreviations: PSA= probabilistic sensitivity analysis; CDC = U.S. Centers for Disease Control and Prevention; Poisson = Poisson distribution; Unif = uniform distribution; Exp = exponential distribution; SEER = Surveillance Epidemiology and End Results program; ADR = adenoma detection rate; Beta = beta distribution; SE = standard error; GI = gastrointestinal; LogN = lognormal distribution; CMS = Centers for Medicare and Medicaid Services.
[a] In multiway probabilistic sensitivity analyses the model parameters were varied randomly according to Uniform, Beta or Lognormal distributions. To limit the degrees of freedom, several parameters were assumed to be perfectly correlated: sensitivity for small adenomas with sensitivity for medium adenomas, sensitivity for large adenomas with sensitivity for cancer, all complication types, costs of colonoscopy with and without polypectomy, and all treatment costs. Other parameters were varied independently.
[b] Sensitivity was defined as the probability of detecting an adenoma that was present at the time of exam. Based on baseline-detection rates in our data, sensitivity for colorectal cancers was assumed to be unrelated to ADR.
[c] The lack of specificity indicates how many of the exams that did not detect adenomatous lesions included polypectomy for non-adenomatous lesions.
[d] Colonoscopy was considered complete if the cecum was reached. In the 2% incomplete examinations, the endpoint was assumed to be distributed uniformly over the colon/rectum.
[e] We assumed that colonoscopy without polypectomy was not associated with a higher risk of complications. The risk of complications for polypectomy was assumed to increase exponentially with age. Serious GI events included perforation and gastrointestinal bleeding requiring blood transfusions; other GI events included paralytic ileus, nausea, vomiting and dehydration, abdominal pain; and cardiovascular events included myocardial infarction or angina, arrhythmias, congestive heart failure, cardiac or respiratory arrest, or syncope, hypotension, or shock. The fatal perforation rate was derived from estimates of the incidence of perforation and case-fatality for perforation.[271 272]
[f] Screen- and treatment costs include patient time costs (opportunity costs of spending time on screening or being treated for a complication or colorectal cancer), but do not include travel costs, costs of lost productivity, and unrelated health care and non-health care costs in added years of life. Patient time was valued at the median US wage in May 2013 ($16.87 per hour), and we assumed that colonoscopies involve 8 hours of patient time. Patient time costs were already included in the estimates for the costs of life-years with cancer care obtained from a study by Yabroff et al.[278]
[g] Care for colorectal cancer was divided in three clinically relevant phases: the initial, continuing, and terminal care phase. 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 cancer patients dying from the disease and cancer 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.
[h] More details regarding the calibrated natural history parameters and other model elements are provided in **Chapter 2**.

In a separate analysis, patient populations in each ADR quintile were simulated using the age distribution at screening (**Appendix 7**). We derived 5 different sets of parameters for per-lesion sensitivity by polyp size to reproduce the average ADR for each quintile. These were constrained by assuming: (1) sensitivity for cancer was 98% across all quintiles; (2) sensitivity for medium to large adenomas varied less than for small adenomas, and increased according to a fixed rule from the lowest

to the highest quintile (fixed detection likelihood (sensitivity/[1-sensitivity]) ratios for adjacent quintiles) while matching estimates for average practice in the middle quintile (85% for medium adenomas, 95% for large adenomas)[163]; (3) maximum sensitivity for adenomas was 98%. Sensitivity for adenomas was then varied to match ADR values with 0.1 point precision. The estimates were independent of adenoma location. KPNC data on cancer diagnoses after colonoscopy were compared to the cancer incidence predicted by the model.

Complication risk of colonoscopy

Adverse events for colonoscopy including polypectomy used age-specific complication rates derived from published literature (**Table 7.1**).[271,272]

Costs of screening and treatment

Approximate societal costs of colonoscopy, complications and colorectal cancer treatment utilized 2007 Medicare payment rates and co-payments (**Table 7.1**) [erratum: treatment cost data represent the period 1998-2003].[152,277,278] All costs included patient time valued at median US wage in 2013, updated to December 2013 based on general Consumer Price Index.[279] Costs of colonoscopy with polypectomy included a variable component for polyp resection and pathology based on number of polyps resected.

Outcomes

Outcomes included were colorectal cancer incidence, mortality, years of life lost, number of colonoscopies, complications, and the costs of screening and treatment in unscreened persons and in those screened according to ADR quintiles. In addition, we estimated the average outcome differences associated with 5 percentage-point higher ADRs using linear regression. Outcomes were discounted to 2010 at a fixed annual rate of 3% and reported with uncertainty ranges.

Analysis

We simulated a US population cohort of 10 million men and women born January 1, 1960. For patients reaching the age of 50 without having colorectal cancer diagnosed (9.4 million), we compared the outcomes of no screening, or of screening colonoscopy at ages 50, 60 and 70 by physicians from one of the five ADR quintiles.[246] Patients with adenomas detected were assumed to receive surveillance according to current United States guidelines.[139] We assumed that the same physician performed all screening and surveillance colonoscopies in each individual patient, and thus, ADR exposure level remained constant during the life-course.

Multiway probabilistic sensitivity analyses were conducted to derive 95% CI's for all outcomes evaluated.[280,281] In 1,000 simulation runs of 10 million persons we varied 13 key parameters along uniform, beta, or lognormal distributions (**Table 7.1**).

Sensitivity analysis

We evaluated the robustness of results using several alternative modeling scenarios. Between-quintile ADR variation was attributed either: entirely to exam sensitivity for small lower-risk adenomas; equally to exam sensitivity for small, medium and large adenomas; or partially to exam sensitivity and to adenoma prevalence or colonoscopy completion rates (~1% higher per percentage-point higher ADR). Adenoma patients received either more intensive or no surveillance. We also evaluated a 50% increased colonoscopy cost level and undiscounted outcomes.

To evaluate data uncertainty, we performed a bootstrap analysis on the association between observed average ADR and interval cancer rates across ADR quintiles and contrasted the resulting weak and strong association samples (2.5-97.5th percentile) to the modeling scenarios.

Statistical Software

For microsimulation modeling we used Delphi 7.0 (Borland Software Corp). Additional data analyses were performed using Stata 13.1 (StataCorp).

RESULTS

A total of 57,588 screening colonoscopies were performed by 136 KPNC physicians during 1998-2010 (**Table 7.2)**. After exclusion of patients with less than 6 months follow-up (n=7,718), there were 179,812 person-years of follow-up time. Interval colorectal cancer incidence per 100,000 person-years varied from 66.6 (95% CI: 43.2-97.0) in ADR quintile 1 to 39.0 (95% CI: 22.7-62.4) in quintile 4, but was 49.7 (95% CI 27.8-81.9) in quintile 5.

Table 7.2 Kaiser Permanente Northern California patient and physician characteristics according to quintile of adenoma detection rate.

Variable	Quintiles of adenoma detection rate					Total
	1	2	3	4	5	
Physician characteristics						
Physicians, n	27	27	28	27	27	136
Adenoma detection rate						
Mean	15.32	21.27	25.61	30.89	38.66	26.45
Median	16.56	21.50	25.70	30.96	38.86	25.70
Range	7.35-19.05	19.06-23.85	23.86-28.40	28.41-33.50	33.51-52.51	7.35-52.51
Patient characteristics						
Screened adults, n	11,799	10,579	10,978	12,918	11,314	57,588
Cancer diagnosed within 6 months	114	93	106	176	119	608
Less than 6 months of follow-up	1,452	1,253	1,179	1,421	1,805	7,110
Proportion male, %	42.8	43.4	44.1	45.0	44.5	44.0
95% CI	(34.6, 51.0)	(36.0, 50.8)	(36.2, 51.9)	(37.3, 52.7)	(37.1, 51.9)	(36.1, 51.8)
Mean age, years	61.3	61.3	62.0	62.0	61.9	61.7
95% CI	(59.3, 63.2)	(59.5, 63.1)	(59.1, 64.9)	(60.1, 64.0)	(59.5, 64.3)	(59.3, 64.1)
Age groups, %						
50-54 years	25.6	25.4	23.5	23.6	24.0	24.4
55-59 years	21.4	20.6	19.9	19.8	19.8	20.3
60-64 years	20.7	21.9	20.7	20.2	20.8	20.8
65-69 years	14.9	15.4	15.0	16.2	15.4	15.4
70-74 years	9.7	9.2	11.4	10.5	10.8	10.3
75-84 years	6.9	6.9	8.9	8.6	8.4	7.9
>85 years	0.8	0.7	0.7	1.0	0.8	0.8
Race/ethnicity, %						
Non-Hispanic white	69.0	73.0	67.9	65.7	66.5	68.3
Hispanic	5.9	5.5	8.2	7.1	8.1	7.0
Non-Hispanic black	7.8	5.3	4.4	4.5	4.0	5.2
Asian	7.4	7.8	10.2	14.5	13.0	10.7
Native Americans	0.3	0.3	0.4	0.4	0.3	0.4
Other	2.3	2.2	2.5	2.5	2.9	2.5
Unknown	7.2	5.8	6.4	5.4	5.2	6.0
Patients with adenomas detected, n [a]	1,808	2,250	2,811	3,991	4,374	15,234
Person-years of follow-up [b]	39,033	33,251	33,564	43,635	30,200	179,682
Interval cancers diagnosed [c]	26	18	14	17	15	90
Incidence per 100,000 yr^{-1}	66.6	54.1	41.7	39.0	49.7	50.1
95% CI	(43.2,97.0)	(32,85.3)	(23.1,70.8)	(22.7,62.4)	(27.8,81.9)	(40.3,61.6)

← **Table 7.2** Kaiser Permanente Northern California patient and physician characteristics according to quintile of adenoma detection rate.
Abbreviations: yr^{-1}= per person-year; CI = confidence interval.
[a] Including only histologically confirmed adenomas by pathologists.
[b] Patients were followed from the date of their index colonoscopy until the first of the following events: negative follow-up colonoscopy, diagnosed cancer, death or departure from membership, 10 years follow-up, or study end (31 December 2010).
[c] Interval cancers were colorectal adenocarcinomas diagnosed ≥ 6 months and ≤ 10 years of the index colonoscopy

Simulated interval cancer incidence

To replicate the average detection rate per ADR quintile in the KPNC cohort in the model, colonoscopy sensitivity was varied according to adenoma size from: 14.7% in quintile 1, 41.0% in quintile 3 to 98% in quintile 5 for small adenomas; 39.6 to 98% for medium adenomas; and 88.0 to 98% for large adenomas (see **Table 7.1** for estimates per ADR quintile). The model closely reproduced observed colorectal cancer incidence in the lower four ADR quintiles, but underestimated incidence in the upper quintile (**Supplementary Figure 7.2**).

Lifetime colorectal cancer outcomes without and with screening

The model estimated average overall life expectancy without exposure to screening and surveillance was 81.1 years, the lifetime colorectal cancer risk was 34.2/1,000 (95%CI:25.9-43.6), lifetime colorectal cancer mortality risk was 13.4/1,000 (95%CI:10.0-17.6), and 138.7 life-years per 1,000 patients (95%CI:103.0-184.0) were lost due to colorectal cancer, which is 10.4 years per cancer death (**Table 7.3**). Among screened patients, simulated lifetime risk of colorectal cancer incidence was on average 19.1/1,000 (95%CI:14.3-24.8), mortality was 3.8/1,000 (95%CI:2.8-5.2); and 42.7 (95%CI:30.9-57.5) life-years per 1,000 patients were lost to the disease.

Table 7.3 Modeling results: Outcomes associated with colonoscopy screening according to quintile of adenoma detection rate. [a b]

Lifetime health	No screening		Screening; Quintiles of adenoma detection rate									
outcomes per			1		2		3		4		5	
1000 patients	Mn	95%CI[c]	Mn	95%CI[c]	Mn	95%CI[c]	Mn	95%CI[c]	Mn	95%CI[c]	Mn	95%CI[c]
CRC outcomes												
CRC cases	34.2	(25.9-43.6)	26.6	(20-34.3)	21.9	(16.3-28.1)	19.0	(14-24.7)	15.6	(11.6-20.3)	12.5	(9.3-16.5)
Advanced cases [d]	16.8	(12.3-22.6)	7.3	(5.2-9.9)	5.6	(4-7.7)	4.7	(3.4-6.3)	3.7	(2.6-5.1)	2.9	(2.1-3.9)
CRC deaths	13.4	(10-17.6)	5.7	(4.2-7.7)	4.5	(3.2-6)	3.7	(2.7-5)	3.0	(2.1-4)	2.3	(1.7-3.1)
Years of life lost [e]	13.7	(103-184)	61.4	(44.4-82.9)	49.2	(35.6-66.1)	41.8	(30.4-55.9)	33.9	(24.4-46.3)	27.0	(19.5-36.2)
Screening effectiveness												
Prevented CRC cases	-	-	7.7	(5.4-10.3)	12.3	(9.1-16.2)	15.3	(11.4-19.8)	18.7	(14-24)	21.7	(16.2-27.8)
Prevented CRC deaths	-	-	7.7	(5.8-10)	8.9	(6.7-11.6)	9.6	(7.2-12.6)	10.4	(7.8-13.8)	11.1	(8.2-14.6)
Years of life saved	-	-	77.3	(58-102.3)	89.5	(66.5-117.1)	96.8	(71.8-127.4)	104.8	(78.2-139.1)	111.7	(82.8-148.4)

Abbreviations: CI = confidence interval; CRC = colorectal cancer.

[a] All outcomes were discounted to 2010 at a fixed rate of 3% per year. For undiscounted outcomes see **Supplementary Table 7.2**.

[b] Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

[c] 95% confidence intervals were derived by multiway probabilistic sensitivity analysis.

[d] Advanced-stage cancers were stage III-IV according to the 5th edition Cancer Staging Manual from the American Joint Committee on Cancer.[149]

[e] Years of life lost to the disease were obtained by subtracting the simulated lifetimes with disease from the simulated lifetimes based on other-cause mortality rates.

The modeled risks were inversely related to the level of adenoma detection (**Table 7.3**). The simulated lifetime risk of colorectal cancer per 1,000 was 26.6 (95%CI:20.0-34.3) for patients of physicians in ADR quintile 1, and was monotonically lower for subsequent quintiles; in ADR quintile 5, the simulated lifetime colorectal cancer risk was 12.5 (95%CI:9.3-16.5). Compared to ADR quintile 1, simulated lifetime risk of colorectal cancer was on average 11.4% (95%CI:10.3-11.9) lower per 5 percentage-point higher ADR (**Figure 7.1**). Similarly, the simulated lifetime risk of colorectal cancer death and associated years-of-life lost per 1,000 patients were lower from quintile 1 (5.7, 95%CI:4.2-7.7 and 61.4, 95%CI:44.4-82.9) to quintile 5 (2.3, 95%CI:1.7-3.1 and 27.0, 95%CI:19.5-36.2). The simulated lifetime risk of colorectal cancer death was on average 12.8% lower (95%CI:11.1-13.7) for every 5 percentage-point higher physician ADR.

Colonoscopy volume and complications

The model's total estimated number of colonoscopies per 1,000 patients was progressively higher from ADR quintile 1 (2,777, 95%CI:2,626-2,943) to quintile 5 (3,376, 95%CI:3,081-3,681) (**Table 7.4**), an average of 4.6% (95%CI:3.6-5.7) for every 5-point higher ADR (**Figure 7.1**). This difference was related to more frequent surveillance in patients of physicians with higher ADR. The simulated lifetime risk (per 1,000) of serious gastrointestinal complications such as post-polypectomy bleeding and perforation was also higher from ADR quintile 1 (2.2, 95%CI:1.5-3.1) to quintile 5 (3.2/1,000, 95%CI:2.3-4.4), as were the overall complications (6.0, 95%CI:4.0-8.5 to 8.9, 95%CI:6.1-12.0) and fatal complications (0.03 to 0.05). Overall, the simulated risk of complications was on average 9.8% (95%CI:7.5-13.2) higher for every 5-point higher ADR.

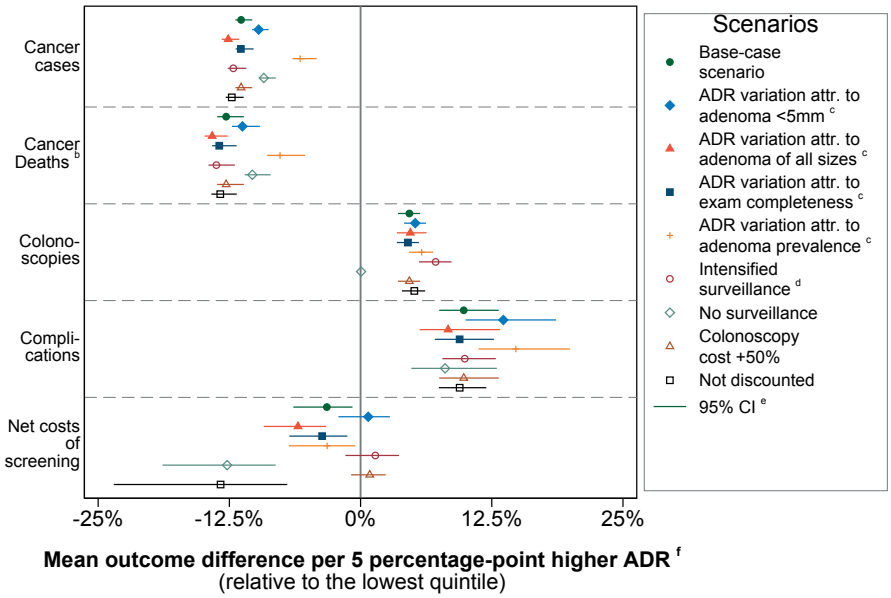


Figure 7.1 Sensitivity analysis results: The adenoma detection rate-outcome relationship for various modeling scenarios. [a]

Abbreviations: ADR = adenoma detection rate; attr. = attributed.

[a] 95% confidence intervals were relatively narrow because we applied the same assumptions for the natural history of colorectal cancer to all patients (**Table 7.1**). Colonoscopy sensitivity was the only assumption varied independently for each ADR quintile.

[b] Results were similar for years of life lost to cancer.

[c] We evaluated four alternative causal models for the observed cancer incidence differences across the ADR quintiles: in scenario 2 all variation in ADR was attributed to sensitivity of colonoscopy for small adenomas under 5 mm, which varied from 5.4 in the lowest quintile to98% in the highest quintile; in scenario 3 all ADR variation was attributed equally to sensitivity for small, medium and large adenomas, which varied from 26.0 to 98%; in scenario 4 it was assumed that the rate of completeness of colonoscopy along with differences in colonoscopy sensitivity accounted for the observed ADR-variations, varying from 75% to 98%; in scenario 4 adenoma prevalence was assumed to be up to a relative 25% higher with higher ADR.

[d] Under intensified surveillance, we assumed that all patients with adenomas detected at colonoscopy underwent surveillance at 3 years after the procedure, and patients with a negative surveillance colonoscopy underwent surveillance at 5 years. For reference, in the base-case analysis, patients with adenomas detected at colonoscopy were referred for surveillance after 3 or 5 years, depending on the number and size of the adenomas detected. Likewise, patients with a negative surveillance colonoscopy were referred for a follow-up colonoscopy in 5 or 10 years, depending on whether the preceding interval was 3 or 5 years.

[e] 95% confidence intervals were derived by multiway probabilistic sensitivity analysis.

[f] The mean differences in simulated outcomes per 5 percentage-point higher ADR were derived by linear regression, and presented relative to the model outcomes for ADR quintile 1 (formula: 5 x $beta_{ols}$ / $outcome_{q1}$).

Estimated costs of screening and treatment

For ADR quintile 1, estimated colonoscopy-related costs in US dollars per 1,000 patients were $2.7 million (95%CI:2.4-3.1), and estimated treatment costs were $2.4 million (95%CI:1.8-3.1), for an estimated total of $5.2 million (95%CI:4.4-6.0) without adjustment and $2.1 million (95%CI:1.8-2.4) with adjustment for the estimated costs without screening (**Table 7.4**). For higher ADR quintiles, estimated colonoscopy costs were higher, but estimated treatment costs were lower, for lower estimated total costs ($4.9 million, 95%CI:4.1-5.6) and net screening costs ($1.8 million, 95%CI:1.3-2.3) in quintile 5. Estimated net screening costs were on average 3.2% lower (95%CI:0.8-6.4) for every 5-point higher ADR.

Sensitivity analyses

The simulations were stable to various assumptions regarding colorectal carcinogenesis, colonoscopy efficacy and surveillance intervals (**Figure 7.1**). Although simulated costs were more unstable, the absolute corresponding cost differences were small (**Supplementary Table 7.1**). Without discounting, the estimated benefits of higher ADR were approximately twice as large as with discounting (**Supplementary Table 7.2-3**).

For ADR quintiles 1 to 4, strong and weak association scenarios from the bootstrap analysis for observed ADR and cancer incidence data were within the predicted ranges of the sensitivity analysis models (**Supplementary Figure 7.2b**).

Table 7.4 Modeling results: Resources and complications for colonoscopy screening according to quintile of ADR. [a b]

Resources and complications	No screening		Screening; Quintiles of adenoma detection rate									
per 1,000 patients			1		2		3		4		5	
	Mn	95% CI[c]	Mn	95% CI[c]	Mn	95% CI[c]	Mn	95% CI[c]	Mn	95% CI[c]	Mn	95%CI[c]
Screening resources used												
Total colonoscopies	-	-	2,777	(2626-2943)	2,980	(2786-3197)	3,094	(2873-3329)	3,252	(2985-3533)	3,376	(3081-3681)
Screening colonoscopies	-	-	2,008	(1972-2041)	1,948	(1901-1993)	1,912	(1858-1964)	1,857	(1794-1921)	1,807	(1736-1885)
Surveillance colonoscopies [d]	-	-	769	(584-968)	1,032	(796-1290)	1,182	(915-1465)	1,395	(1074-1728)	1,569	(1204-1935)
Colonoscopies with polypectomy (screening and surveillance)	-	-	956	(742-1176)	1,187	(938-1424)	1,312	(1045-1553)	1,479	(1188-1733)	1,599	(1284-1862)
Colonoscopy-related complications	-	-	6.0	(4-8.5)	7.4	(5-10.1)	8.0	(5.4-10.8)	8.6	(5.9-11.7)	8.9	(6.1-12)
Serious GI complications	-	-	2.2	(1.5-3.1)	2.7	(1.8-3.7)	2.9	(2-4)	3.2	(2.2-4.3)	3.2	(2.3-4.4)
Fatal GI complications [e]	-	-	0.03	n.a.	0.04	n.a.	0.04	n.a.	0.05	n.a.	0.05	n.a.
Other GI complications	-	-	2.2	(1.4-3.1)	2.6	(1.8-3.6)	2.8	(2-3.9)	3.1	(2.1-4.2)	3.2	(2.2-4.3)
Cardiovascular complications	-	-	1.6	(1.1-2.3)	2.0	(1.4-2.7)	2.1	(1.5-2.9)	2.3	(1.6-3.2)	2.4	(1.7-3.2)
Financial resources used (US $ million) [f]												
Total medical costs	3.1	(2.3-4)	5.2	(4.4-6)	5.1	(4.3-5.9)	5.0	(4.2-5.8)	4.9	(4.2-5.7)	4.9	(4.1-5.6)
Screening costs	-	-	2.8	(2.5-3.1)	3.1	(2.7-3.4)	3.2	(2.8-3.7)	3.5	(3-4)	3.7	(3.2-4.2)
Colonoscopy costs	-	-	2.7	(2.4-3.1)	3.0	(2.6-3.4)	3.2	(2.8-3.6)	3.4	(3-3.9)	3.6	(3.1-4.2)
Complication costs	-	-	0.0	(0-0.1)	0.0	(0-0.1)	0.0	(0-0.1)	0.1	(0-0.1)	0.1	(0-0.1)
Treatment costs	3.1	(2.3-4)	2.4	(1.8-3.1)	2.0	(1.5-2.6)	1.7	(1.3-2.3)	1.5	(1.1-1.9)	1.2	(0.9-1.5)
Net screening costs [g]	-	-	2.1	(1.8-2.4)	2.0	(1.6-2.4)	1.9	(1.5-2.3)	1.9	(1.4-2.3)	1.8	(1.3-2.3)

Abbreviations: CI = confidence interval; GI = gastrointestinal; n.a. = not assessed.

[a] All outcomes were discounted to 2010 at a fixed rate of 3% per year. For undiscounted outcomes see **Supplementary Table 7.3**.

[b] Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

[c] 95% confidence intervals were derived by multiway probabilistic sensitivity analysis.

[d] Patients with adenomas detected had surveillance colonoscopies 3 years after the detection of ≥1 large adenoma or ≥3 adenomas of any size, and 5 years after the detection of ≤3 adenomas with a diameter of <10 mm. In case of a negative surveillance colonoscopy, the next interval was 5 or 10 years, depending on whether the length of the preceding interval was 3 or 5 years. Surveillance was continued until death or diagnosis of cancer.

[e] For the simulated effect of fatal perforations on life-years lost, we assumed immediate death.

[f] Besides resources for endoscopy and endoscopy-related complications, screening colonoscopy also influenced the modeled resources for cancer care. Higher ADR was associated in the model with a lower use of these resources, because of lower associated cancer incidence.

[g] Net screening costs were derived by comparing the estimated total medical costs in case of screening (screening and cancer treatment costs) with the total medical costs in case of no screening. Minor inconsistencies in the resulting net costs are due to rounding.

DISCUSSION

This study used data from a large community-based United States healthcare system in a microsimulation model to estimate the lifetime outcomes and costs of colonoscopy screening at different levels of adenoma detection.[124] Our results suggest that higher adenoma detection rates may be associated with up to 50-60% lower lifetime colorectal cancer incidence and mortality without higher net screening costs despite a higher number of colonoscopies and polypectomy-associated complications.

The model's differences in observed interval colorectal cancer incidence were assumed to result from differences in the sensitivity of the exam, particularly for small-to-medium-sized adenomas. However, ADR may act as a surrogate for other aspects of colonoscopy quality, such as the test completeness, adequacy of lesion resection, and removal of more aggressive lesions such as sessile serrated polyps.[282] Although some of these alternative explanations were evaluated in sensitivity analyses, with similar long-term results (**Figure 7.1**), we could not establish which factors accounted for the observed differences (**Supplementary Figure 7.2b**), and whether others might be involved.

The frequency and intensity of surveillance of adenoma patients may also contribute to patient outcome differences, because higher ADRs increase the number of patients for active surveillance.[139] However, sensitivity analyses indicated that surveillance did not account for the simulated survival benefits for patients of physicians with higher ADRs (**Figure 7.1**). Future research is needed to assess whether the current intensity of surveillance is still appropriate if test sensitivity further increases.

Prior studies have shown an inverse relationship between ADR level and the patient's risk of colorectal cancer up to 5 years after colonoscopy.[125,127,283] A recent large study found that patients of physicians in the highest ADR quintile had a 48% lower disease risk and a 62% lower mortality risk compared to the lowest quintile.[124] Adenoma detection rates may relate to patient outcomes over a lifetime of colonoscopy screening and surveillance. Our model estimated that discounted lifetime incidence and mortality risks averaged 11-13% lower for every 5-point higher ADR, which translates to overall differences of 53-60% between the lowest and highest ADR quintiles. Higher ADR was associated in the model with up to 34.4 additional life-years saved per 1,000 patients, which represents about 10 years per prevented cancer death, 2 weeks per average patient, and one-third of the maximum potential mortality benefit derived from screening (5 weeks per patient).

Screening colonoscopy is considered cost-effective for preventing colorectal cancer through adenoma detection and removal.[102,246] However, it has been suggested that incentivizing higher adenoma detection, for example through value-based purchasing programs,[284] could lead to unacceptably higher cost because of more

frequent surveillance in patients with low-risk adenomas.[141] Our model suggests that higher detection rates are associated with only a moderately higher total number of colonoscopies: although the average surveillance patient in the modeling analysis received about twice as many procedures as a patient without detected adenomas, the additional proportion of patients undergoing surveillance with higher detection rates was limited to a maximum of 17%. By evaluating the costs for screening, surveillance, screening-associated complications and cancer care, our model suggested that ADR is not associated with higher overall costs.

Another theoretical disadvantage of higher ADRs is a higher risk of complications due to more colonoscopies and polypectomies. The model suggested that for every 5-point higher ADR the lifetime complication risk is on average 10% higher. The corresponding absolute risk difference of 0.6/1,000 was counterbalanced in the model by a 3.0/1,000 lower risk of colorectal cancer and a 0.7/1,000 lower risk of disease-related mortality (**Supplementary Table 7.1**). Our model included mild gastrointestinal symptoms such as nausea or abdominal pain and rare fatal complications. The model's complication rates are somewhat lower than those presented by other studies,[282] because we adjusted our estimates for the risk of similar events in the group unexposed to colonoscopy.[271]

The model predicted all colorectal cancer outcomes to be lower for every higher quintile of adenoma detection. These predictions closely replicated the observed interval cancer incidence in the lower four ADR quintiles, but underestimated adenoma detection and interval cancer incidence for the highest quintile (**Supplementary Figures 7.2-3**). Although this suggests more uncertainty for the associations beyond approximately 30% (quintile 4 average), in a much larger sample of colonoscopies for all indications from the same data source, a plateau in outcome differences across ADR quintiles was not observed.[124]

This study has some other potential limitations. First, we confined the ADR estimates and analyses to screening colonoscopies. This decreased the number of interval cancers and therefore the precision. However, sensitivity analyses indicated that this did not have a strong effect on long-term model projections (**Supplementary Figure 7.2b**). Second, modeled colorectal adenomas and cancer risk without screening included >10-year-old data. Uncertainty in corresponding model parameters was assessed with probabilistic sensitivity analysis. Third, our findings for the average association between ADR and patient outcomes do not necessarily mean that modifying ADR alone in individual physicians would lead to fewer interval cancers for their patients, given modeling cannot prove causal relationships. Fourth, our estimates assumed compliance with screening and surveillance guidelines and that patients receive colonoscopies from physicians with similar ADRs throughout their lifetimes. Finally, our cost estimates used Medicare rates and co-payments without

supplemental anesthesia costs, and thus may not represent true societal screening costs.[285] We also assumed that there was no overuse of surveillance or screening.[62] However, sensitivity analyses suggested that these surveillance and cost-related factors may not have a large net effect (**Figure 7.1**, **Supplementary Table 7.1**).

Conclusions

In this microsimulation modeling study, higher adenoma detection rates in screening colonoscopy were associated with lower lifetime risks of colorectal cancer incidence and mortality without being associated with substantially higher overall costs. Future research including other direct colonoscopy quality indicators is needed to assess why adenoma detection rates vary, and if increasing adenoma detection would be associated with improved patient outcomes.

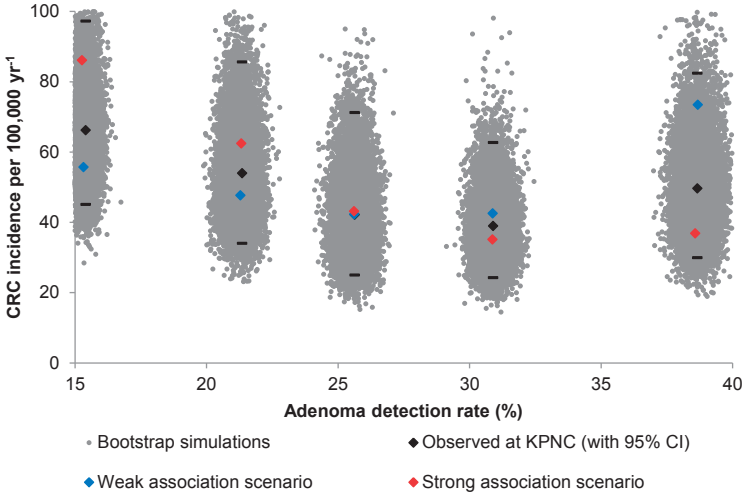
APPENDIX 7

To derive point estimates of per-lesion sensitivity of colonoscopy for each quintile of adenoma detection in the Kaiser Permanente Northern California (KPNC) data, we simulated the screened populations in each quintile in terms of the age distribution at the time of screening. Population size was inflated in the model to 1 million lives per adenoma detection rate (ADR) quintile to reduce random variability in model outcomes. Two main simplifications were that: (1) although we simulated the age distribution of patients per ADR quintile, inter-provider differences in terms of patient risk factors such as age and sex were assumed to be negligible. Thus, apart from the different age distributions per ADR quintile, all simulated patients were selected randomly from an average-risk US population; (2) it was assumed that patients did not get screened previously, whereas the data included some individuals with a negative prior colorectal cancer test (≥10 years ago). Any misclassification was assumed to be non-differential given random assignment to each ADR quintile.

To validate the model including the point estimates for colonoscopy sensitivity in terms of the predicted interval cancer incidence after screening, we also simulated the follow-up time as included in the KPNC data. Because the incidence rate is variable over time and depends on whether a person had adenomas detected at baseline, we exactly replicated the person-years of follow-up after 1, 2, …, 10 years, stratifying patients with a positive and negative baseline colonoscopy (for adenomas). Because the interval cancers in the data included cancers detected by opportunistic screening or surveillance colonoscopies, we also simulated the proportion of patients with a repeat colonoscopy in years 1, 2, …, 10.

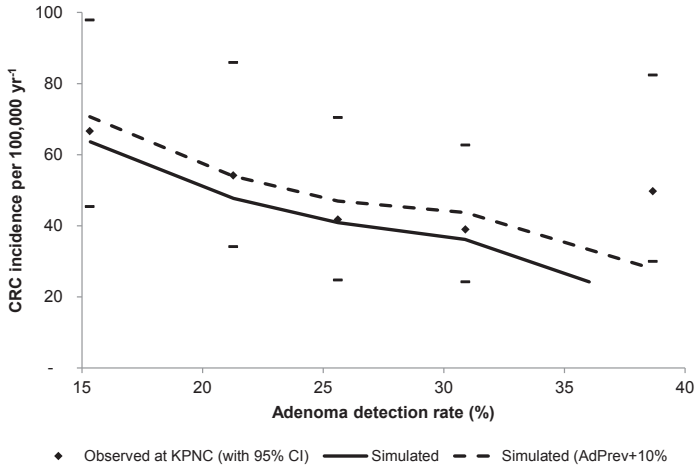
The 5 different sets of parameters for per-lesion sensitivity by polyp size were derived to reproduce the average ADR for each quintile. The parameters were constrained by assuming that: (1) sensitivity for cancer was 98% across all quintiles; (2) sensitivity for medium to large adenomas varied less than for small adenomas, and increased according to a fixed rule from the lowest to the highest quintile (fixed detection likelihood (sensitivity/[1-sensitivity]) ratios for adjacent quintiles) while matching estimates for average practice in the middle quintile (85% for medium adenomas, 95% for large adenomas);[163] (3) maximum sensitivity for adenomas was 98%. Sensitivity for adenomas was then varied to match ADR values with 0.1 point precision. The estimates were independent of adenoma location. From the lowest to the highest ADR quintile, resultant sensitivity was 14.7% in quintile 1, 41.0% in quintile 3 and 98% in quintile 5 for small adenomas, 39.6 to 98% for medium adenomas, and 88.0 to 98% for large adenomas (see **Table 7.2** for estimates per ADR quintile).

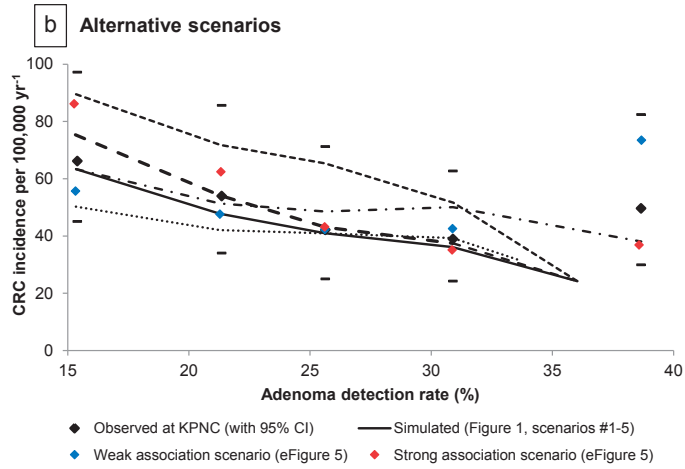
KPNC data on cancer diagnoses after colonoscopy were compared to the cancer incidence predicted by the model. The model closely reproduced observed incidence in the lower four ADR quintiles, but underestimated incidence in the upper quintile (**Supplementary Figure 7.2-3**).



Supplementary Figure 7.1 Bootstrap analysis for average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California. [a]

[a] We performed a parametric bootstrap analysis for average observed adenoma detection and incidence rates per ADR quintile (100,000 scenarios, 10,000 shown). Incidence was varied along the lognormal distribution (with Poisson standard errors) and adenoma detection was varied along the normal distribution (with binomial standard errors). Weak and strong association scenarios represent the resulting 2.5th (average 2.4-6th) and 97.5th (97.4-6th) percentile of bootstrap scenarios in terms of the linear regression coefficient for incidence to ADR.



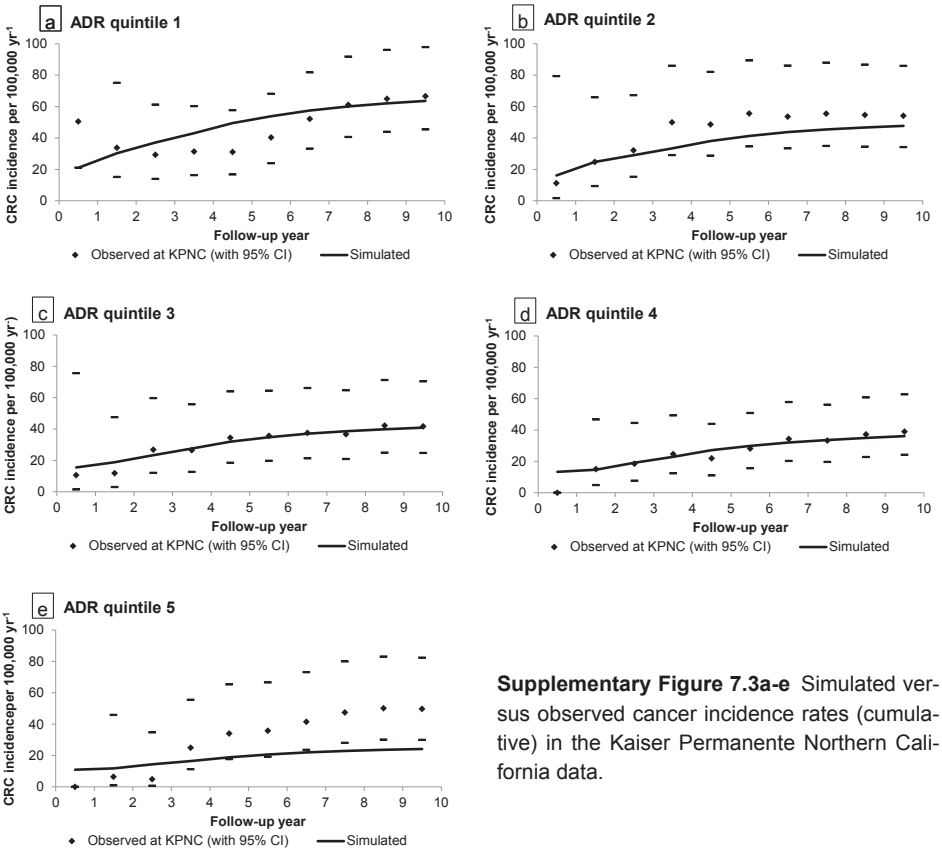


Supplementary Figure 7.2a-b Simulated versus observed average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California. [a]

Abbreviations: Ad. = Adenoma; prev. = prevalence; assoc . = association.

***** In the base-case model the adenoma prevalence of 37% was insufficient to reproduce ADR levels observed for the upper quintile in the KPNC data (the curve stops below 37% due to imperfect sensitivity). The dashed scenario in panel a with higher simulated adenoma prevalence reproduced the observed ADR level and led to similar overall results as the base-case model (*not shown*).

[a] In panel b, supplementary eFigure 5 in the legend corresponds to **Supplementary Figure 7.1** in this thesis, and Figure 1 corresponds with **Figure 7.1** in this thesis.



Supplementary Figure 7.3a-e Simulated versus observed cancer incidence rates (cumulative) in the Kaiser Permanente Northern California data.

Supplementary Table 7.1 Sensitivity analysis results: The adenoma detection rate-outcome relationship for various modeling scenarios.

Scenario		Average outcome difference per 5 percentage-point higher adenoma detection rate [a]									
		Cancer cases		Cancer deaths		Colonoscopies		Complications		Net screening cost	
		Mean	95%CI	Mean	95%CI	Mean	95%CI	Mean	95%CI	Mean	95%CI
1. Base-case	Rel., %	-11.4	(-11.9,-10.3)	-12.8	(-13.7,-11.1)	4.6	(3.6,5.7)	9.8	(7.5,13.2)	-3.2	(-6.4,-0.8)
	Abs.	-3.0	(-3.9,-2.2)	-0.7	(-1,-0.5)	129.1	(94.8,163.4)	0.6	(0.4,0.8)	-0.1	(-0.1,0)
2. ADR-variation attr. to adenoma ≤5mm [b]	Rel., %	-9.7	(-10.3,-8.8)	-11.3	(-12.3,-9.6)	5.2	(4.1,6.2)	13.6	(10,18.6)	0.7	(-2.1,2.8)
	Abs.	-2.3	(-3,-1.7)	-0.6	(-0.8,-0.4)	141.2	(106.9,176.9)	0.7	(0.5,1)	0.0	(0,0.1)
3. ADR-variation attr. to adenoma of all sizes [b]	Rel., %	-12.6	(-13.2,-11.5)	-14.1	(-14.9,-12.7)	4.7	(3.5,6.3)	8.3	(5.6,13.3)	-6.0	(-9.2,-3.3)
	Abs.	-3.8	(-5,-2.8)	-1.0	(-1.3,-0.7)	131.0	(93.9,175.7)	0.5	(0.3,0.8)	-0.1	(-0.2,-0.1)
4. ADR-variation attr. to exam completion [b]	Rel., %	-11.4	(-11.9,-10.2)	-13.5	(-14.1,-11.8)	4.5	(3.5,5.6)	9.4	(7.1,12.7)	-3.7	(-6.8,-1.2)
	Abs.	-3.1	(-4,-2.3)	-0.8	(-1.1,-0.6)	126.0	(91.9,159.9)	0.6	(0.4,0.8)	-0.1	(-0.1,0)
5. ADR-variation attr. to adenoma prevalence [b]	Rel., %	-5.8	(-6.5,-4.2)	-7.7	(-8.9,-5.3)	5.8	(4.6,6.9)	14.8	(11.2,19.9)	-3.2	(-6.8,-0.5)
	Abs.	-1.5	(-2.1,-1)	-0.4	(-0.6,-0.3)	161.5	(122.4,199.3)	0.9	(0.6,1.2)	-0.1	(-0.1,0)
6. Intensified surveillance [c]	Rel., %	-12.1	(-12.7,-10.9)	-13.7	(-14.5,-12)	7.2	(5.6,8.7)	9.9	(7.8,12.9)	1.4	(-1.4,3.7)
	Abs.	-3.1	(-4,-2.3)	-0.7	(-1,-0.5)	222.3	(162.5,280.9)	0.7	(0.4,0.9)	0.0	(0,0.1)
7. No surveillance	Rel., %	-9.2	(-9.7,-8.1)	-10.3	(-11,-8.6)	0.0	(0,0.1)	8.1	(4.8,13)	c.s.	(-18.9,-8.1)
	Abs.	-2.6	(-3.4,-1.9)	-0.7	(-1,-0.5)	1.0	(0.7,1.3)	0.3	(0.2,0.4)	-0.2	(-0.3,-0.1)
8. Colonoscopy costs +50%	Rel., %	-11.4	(-11.9,-10.3)	-12.8	(-13.7,-11.1)	4.6	(3.6,5.7)	9.8	(7.5,13.2)	0.9	(-0.9,2.4)
	Abs.	-3.0	(-3.9,-2.2)	-0.7	(-1,-0.5)	129.1	(94.8,163.4)	0.6	(0.4,0.8)	0.0	(0,0.1)
9. No discounting	Rel., %	-12.3	(-12.9,-11.1)	-13.4	(-14.2,-11.8)	5.1	(3.9,6.2)	9.4	(7.5,12)	-13.3	(-23.5,-7)
	Abs.	-5.9	(-7.6,-4.3)	-1.6	(-2.1,-1.1)	192.2	(140,244.2)	1.0	(0.7,1.4)	-0.3	(-0.4,-0.2)

Abbreviations: ADR = adenoma detection rate; CI = confidence interval; Rel. = relative; Abs. = Absolute; attr. = attributable.

[a] Relative outcomes differences were estimated by linear regression and are compared to the lower ADR quintile. Absolute differences are presented as risk/number per 1,000 adults. Absolute cost differences are in US $ million.

[b] We evaluated four alternative causal models for the observed ADR differences across the quintiles: in scenario 2 all variation in ADR was attributed to sensitivity of colonoscopy for small adenomas under 5 mm, which varied from 5.4 in the lowest quintile to98% in the highest quintile; in scenario 3 all ADR variation was attributed equally to sensitivity for small, medium and large adenomas, which varied from 26.0 to 98%; in scenario 4 it was assumed that the rate of completeness of colonoscopy along with differences in colonoscopy sensitivity accounted for the observed ADR-variations, varying from 75% to 98%; in scenario 4 adenoma prevalence was assumed to be up to a relative 25% higher with higher ADR.

[c] Under intensified surveillance, we assumed that all patients with adenomas detected at colonoscopy underwent surveillance at 3 years after the procedure, and patients with a negative surveillance colonoscopy underwent surveillance at 5 years. For reference, in the base-case analysis, patients with adenomas detected at colonoscopy were referred for surveillance after 3 or 5 years, depending on the number and size of the adenomas detected. Likewise, patients with a negative surveillance colonoscopy were referred for a follow-up colonoscopy in 5 or 10 years, depending on whether the preceding interval was 3 or 5 years.

Supplementary Table 7.2 Modeled results: Effectiveness of screening colonoscopy according to quintile of adenoma detection rate (0% discounted). [a b]

Lifetime health outcomes per	No screening		Screening; Quintiles of adenoma detection rate									
1,000 patients			1		2		3		4		5	
	Mn	95%CI	Mn	95%CI	Mn	95%CI	Mn	95%CI	Mn	95%CI	Mn	95%CI
CRC outcomes												
CRC cases	66.8	(50.7-85.1)	48.1	(36.1-62.2)	38.8	(28.9-49.9)	32.9	(24.4-42.9)	26.4	(19.7-34.7)	20.6	(15.4-27.1)
Advanced cases	32.5	(23.7-43.6)	13.4	(9.6-18.4)	10.2	(7.2-13.8)	8.3	(6-11.2)	6.4	(4.5-8.9)	4.9	(3.4-6.6)
CRC deaths	27.8	(20.8-36.5)	11.8	(8.6-15.8)	9.0	(6.5-12)	7.5	(5.4-9.9)	5.8	(4.2-7.9)	4.4	(3.2-5.9)
Years of life lost	324.6	(242-429.1)	141.6	(102.5-190.9)	112.1	(80.9-149.9)	94.6	(68.9-125.9)	75.8	(54.5-103)	59.6	(43.1-79.9)
Screening effectiveness												
Prevented CRC cases	-		18.7	(13.7-24.5)	28.0	(21.3-36.3)	33.8	(25.6-43.4)	40.3	(30.5-51.2)	46.1	(34.6-58.6)
Prevented CRC deaths	-		16.0	(12.1-20.9)	18.8	(14.1-24.3)	20.3	(15.3-26.5)	22.0	(16.6-28.9)	23.4	(17.4-30.7)
Years of life saved	-		183.1	(137.7-241.1)	212.5	(158.2-277.1)	230.0	(171.2-301.6)	248.9	(186.1-329.7)	265.0	(196.9-350.4)

Abbreviations: CI = confidence interval.

[a] Unlike the base-case results in **Table 7.3**, these results were not discounted and represent actual expected lifetime benefits of colonoscopy screening.

[b] Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

Supplementary Table 7.3 Modeled results: Resources and complications for colonoscopy screening according to quintile of adenoma detection rate (0% discounted). [a b]

Resources per 1,000 patients	No screening		Screening; Quintiles of adenoma detection rate									
			1		2		3		4		5	
	Mean	95%CI	Mean	95%CI	Mean	95%CI	Mean	95%CI	Mean	95%CI	Mean	95%CI
Screening resources used												
Total colonoscopies	-	-	3,756	(3505-4029)	4,045	(3729-4388)	4,211	(3849-4594)	4,453	(4021-4900)	4,645	(4151-5135)
Screening colonoscopies	-	-	2,526	(2469-2579)	2,433	(2357-2504)	2,376	(2292-2460)	2,293	(2194-2395)	2,219	(2108-2340)
Surveillance colonoscopies	-	-	1,229	(932-1548)	1,612	(1231-2027)	1,835	(1404-2288)	2,160	(1644-2704)	2,426	(1841-3015)
Colonoscopies with polypectomy (screening and surveillance)	-	-	1,309	(1028-1574)	1,601	(1276-1894)	1,759	(1407-2064)	1,976	(1601-2305)	2,132	(1728-2478)
Colonoscopy-related complications	-	-	10.9	(7.3-15.4)	13.4	(9.2-18.5)	14.5	(9.9-19.7)	15.5	(10.7-21.2)	15.8	(10.8-21.5)
Serious GI complications	-	-	4.0	(2.7-5.6)	4.9	(3.4-6.8)	5.3	(3.6-7.2)	5.7	(3.9-7.8)	5.8	(4-7.9)
Fatal GI complications	-	-	0.04	n.a.	0.05	n.a.	0.06	n.a.	0.06	n.a.	0.07	n.a.
Other GI complications	-	-	3.8	(2.6-5.5)	4.7	(3.3-6.5)	5.1	(3.5-7)	5.5	(3.8-7.5)	5.6	(3.8-7.6)
Cardiovascular complications	-	-	3.0	(2-4.3)	3.7	(2.6-5.2)	4.0	(2.8-5.5)	4.3	(3-5.9)	4.4	(3-6)
Financial resources used (US $ million)												
Total medical costs	6.5	(4.9-8.3)	8.6	(7.1-10.3)	8.1	(6.8-9.6)	7.8	(6.5-9.2)	7.5	(6.3-8.9)	7.3	(6.1-8.5)
Screening costs	-	-	3.8	(3.3-4.2)	4.2	(3.6-4.7)	4.4	(3.8-5)	4.8	(4.1-5.5)	5.1	(4.3-5.8)
Colonoscopy costs	-	-	3.7	(3.2-4.2)	4.1	(3.6-4.7)	4.3	(3.7-4.9)	4.7	(4-5.4)	5.0	(4.2-5.7)
Complication costs	-	-	0.1	(0-0.1)	0.1	(0.1-0.1)	0.1	(0.1-0.1)	0.1	(0.1-0.1)	0.1	(0.1-0.1)
Treatment costs	6.5	(4.9-8.3)	4.9	(3.6-6.3)	4.0	(2.9-5.2)	3.4	(2.5-4.5)	2.8	(2-3.7)	2.2	(1.6-2.9)
Net screening costs	-	-	2.1	(1.6-2.6)	1.7	(1-2.3)	1.3	(0.5-2)	1.1	(0.2-1.9)	0.8	(-0.2-1.7)

Abbreviations: GI = gastrointestinal; CI = confidence interval; n.a. = not assessed.

[a] Unlike the base-case results in **Table 7.4**, these results were not discounted and represent actual expected lifetime resources used in colonoscopy screening.

[b] Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

Chapter 8

Impact of adenoma detection rates on the benefits of fecal testing versus colonoscopy for colorectal cancer

Submitted:

RG Meester, CA Doubeni, A Zauber et al.

ABSTRACT

OBJECTIVE: To estimate how variation in colonoscopy quality, as measured by adenoma detection rate (ADR), influences the benefits of fecal immunochemical testing (FIT) compared with primary colonoscopy screening for colorectal cancer.

DESIGN: Using an established microsimulation model, we estimated the benefits of annual FIT screening at differing ADR levels (quintiles; averages 15.3-38.7%), with colonoscopy screening as comparator. Assumptions used community-based data on physician ADRs and patient's post-colonoscopy risk of cancer. Primary study outcomes were simulated lifetime colorectal cancer incidence and mortality per 1000 patients with probability intervals (PI) from probabilistic sensitivity analysis.

RESULTS: For patients receiving FIT screening with potential follow-up colonoscopy by physicians from the highest ADR quintile, simulated lifetime cancer incidence and mortality were 28.8 (95%PI, 19.8-42.6) and 5.4 (95%PI, 3.5-8.4) per 1000, respectively, versus 20.6 (95%PI, 15.4-27.1) and 4.4 (95%PI, 3.2-5.9) for primary colonoscopy screening (risk ratios, RR=1.40; 95%PI, 1.09-1.89, and RR=1.22; 95%PI, 0.92-1.75). With every 5% point ADR decrease, lifetime cancer incidence was estimated to increase on average 8.6% (95%PI, 5.5-11.4) for FIT versus 12.3% (95%PI, 11.1-12.9) for colonoscopy, and mortality increased 9.4% (95%PI, 6.0-12.7) and 13.3% (95%PI, 11.8-14.2), respectively. In ADR quintile 1, simulated mortality was lower for FIT than colonoscopy screening (10.1; 95%PI, 7.3-13.5, versus 11.8; 95%PI, 8.6-15.8, RR=0.85; 95%PI, 0.82-0.93), while incidences were more similar.

CONCLUSION: Relative cancer incidence and mortality reductions for FIT versus colonoscopy screening may differ by ADR. There may be fewer deaths for colonoscopy screening in higher ADR settings and fewer deaths for FIT in lower ADR settings.

INTRODUCTION

Colorectal cancer is a leading cause of cancer deaths that is largely preventable through screening.[103,197] Colonoscopy is indispensable for colorectal cancer screening, as either a primary screening test or for diagnostic follow-up of positive tests results from other screening methods. Colonoscopy quality, as measured by adenoma detection rate (ADR), or the proportion of a physician's screening exams detecting adenomas, varies widely across providers. ADR has been shown to be inversely related to subsequent cancer incidence and mortality risks among patients undergoing screening colonoscopy.[124,263]

Annual fecal immunochemical testing (FIT) is increasingly used as either a primary colorectal cancer screening method or as an adjunct to colonoscopy-based screening programs to increase overall population screening rates.[4] FIT and colonoscopy screening strategies each have their advantages and disadvantages. Colonoscopy screening is more sensitive for cancers and adenomas and has a long screening interval. FIT may be more acceptable to patients because of the lack of dietary restrictions, the non-invasive nature, and lower risk.[86] Although FIT screening requires diagnostic colonoscopy follow-up of positive results, the overall effectiveness of FIT-based screening may also be affected less by lower ADR levels than primary colonoscopy screening given FIT primarily detects more advanced lesions.[161] However, currently no data exist to compare the benefits of colonoscopy and FIT screening at different ADR levels.

The purpose of this study is to use microsimulation modeling with community-based data,[124] to compare the benefits of a program of annual FIT versus colonoscopy every ten years at various ADR levels.

METHODS

This study used the Microsimulation Screening Analysis-Colon (MISCAN-Colon) model, developed by the Erasmus MC University Medical Center, Rotterdam, Netherlands (**Chapter 2**). The model, its main assumptions, and results for colonoscopy screening have been published.[263]

Test performance assumptions

In this study, assumed variation in colonoscopy performance was based on previously published data from Kaiser Permanente Northern California, an integrated healthcare delivery system in the United States with a well-defined denominator population.[124,263] In Corley et al, ADR quintile averages (ranges) varied: 15.3% (7.35-19.05) for quintile 1; 21.3% (19.06-23.85) for quintile 2; 25.6% (23.86-28.40) for quintile 3; 30.9% (28.41-33.50) for quintile 4; and 38.7% (33.51-52.51) for quintile

5.[124] Corresponding estimates of per-lesion sensitivity of colonoscopy were estimated to vary from quintile 1-5: 14.7-98% for adenomas of 0-5mm in diameter, 39.6-98% for adenomas of 6-9mm, and 88.0-98% for adenomas of ≥10mm (**Table 8.1**).[263] The assumed rate of colonoscopy completeness was fixed at 98% for all ADR quintiles.

Table 8.1. Test performance assumptions in MISCAN

Performance characteristic, %	Colonoscopy, by quintile [a] (screening, diagnostic, surveillance)					FIT
	1	2	3	4	5	
Sensitivity per lesion [b]						
Adenomas ≤ 5 mm	14.7	29.6	41.0	66.2	98	0.0
Adenomas 6 - 9 mm	39.6	65.8	85	94.3	98	11.4 [2.5][e]
Adenomas ≥ 10 mm	88.0	92.2	95	96.8	98	15.9 [3.5][e]
Stage I-IV cancer	98	98	98	98	98	63/89 [3.5][e]
Specificity [c]	100	100	100	100	100	96.4 [2.5][e]
Completeness colonoscopy [d]	98	98	98	98	98	-

Abbreviations: FIT = fecal immunochemical test; ADR q*i* = adenoma detection rate quintile *i*.

[a] Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.[124]

[b] The adenoma sensitivity estimates for FIT (OC Sensor, cutoff >20 µg/g) were obtained by calibrating our model outcomes to the estimated per-person sensitivities from Imperiale et al.[161] The per-person sensitivity of FIT for adenomas, advanced adenomas, and cancer was 7.6, 23.8, 73.8, respectively. We assumed that fecal occult blood testing is more sensitive in cancers towards the end of the occult invasive period (close, time-wise, to becoming symptomatic): for preclinical cancers which will become symptomatic within the same stage, assumed test sensitivity was higher. This assumption showed good concordance with guaiac fecal occult blood test trial results.[190] Colonoscopy sensitivity estimates were derived elsewhere.[263]

[c] The probability of a false positive result was random in the base-case analysis, and independent of person or lesion. We assumed perfect specificity for colonoscopy including pathological examination of detected lesions.

[d] This is the proportion of colonoscopies visualizing the maximum point of reach of the endoscope, i.e. the cecum.

[e] Standard deviation for the probabilistic sensitivity analysis is shown in brackets. A Beta distribution was assumed to reflect uncertainty.

The modeled effectiveness of FIT-based screening (OC Sensor test with a positivity cutoff of 20 µg/g cutoff) is based both on the sensitivity and specificity of FIT and the sensitivity and completeness of the colonoscopy exam used for follow-up of positive FIT results. Colonoscopy performance assumptions were varied as above for colonoscopy screening according to ADR level. Assumed per-lesion sensitivity of FIT was derived from recently published observational data, and was 4.9% for adenomas of 6-9mm, 16.2% for adenomas ≥10mm, and 64-89% for cancer (**Table 8.1**).[161]

Analysis

For this study, MISCAN-Colon was used to generate an average-risk screening population of ten million men and women born on January 1, 1965. Patients received annual FIT between the ages 50-75 years.[246] Patients with a positive FIT received follow-up colonoscopy. Patients with adenomas detected in screening received colonoscopy surveillance according to the most current guidelines.[139] We compared colorectal cancer outcomes for FIT according to level of adenoma detection. For reference, we also estimated outcomes with colonoscopy screening and without any screening.

Primary study outcomes were simulated lifetime colorectal cancer incidence and mortality according to ADR quintile (undiscounted). We also estimated the continuous change in outcomes per 5% point increase in ADR using linear regression. Multivariate probabilistic sensitivity analysis was used to derive 95% probability intervals (95% CI) for all model outcomes. In 1000 simulation runs of 10 million persons we varied 13 key parameters along uniform, beta, or lognormal distributions.[263] FIT performance assumptions were also varied (**Table 8.1**).

We conducted one-way sensitivity analyses repeating our estimation of the continuous change in outcomes for every 5% point lower ADR, assuming 5-15% point lower or higher FIT sensitivity, 2.5% point lower or higher FIT specificity, and also varying the extent to which ADR variation was attributed by the model to colonoscopy sensitivity for diminutive lesions. In the one extreme, all ADR variation was attributed to variation in small adenoma miss rates, in the other extreme, physicians were assumed to miss all sizes of lesions with equal probability.

Funding

MISCAN-colon is part of the Cancer Intervention and Surveillance Modeling Network (CISNET) sponsored by the United States National Cancer Institute (NCI). This work was supported by the NCI-funded consortium Population-Based Research Optimizing Screening through Personalized Regimens (PROSPR), the overall aim of which is to conduct multi-site, coordinated, transdisciplinary research to evaluate and improve cancer screening processes.

RESULTS

Among unscreened patients, the simulated lifetime risk of colorectal cancer was 66.8 (95%PI, 50.7-85.1) per 1000, and the simulated risk of colorectal cancer mortality was 27.8 (95%PI, 20.8-36.5) per 1000 (**Figure 8.1**). Among patients screened with colonoscopy, the average simulated colorectal cancer incidence and mortality risks

across all ADR quintiles were 33.4 (95%PI, 24.8-42.8) and 7.7 (95%PI, 5.6-10.2) per 1000, respectively. Among patients screened with FIT (with colonoscopy follow-up for positive results), the average simulated colorectal cancer incidence and mortality risks were 37.9 (95%PI, 27.8-52.3) and 7.4 (95%PI, 5.3-10.3) per 1000, respectively.

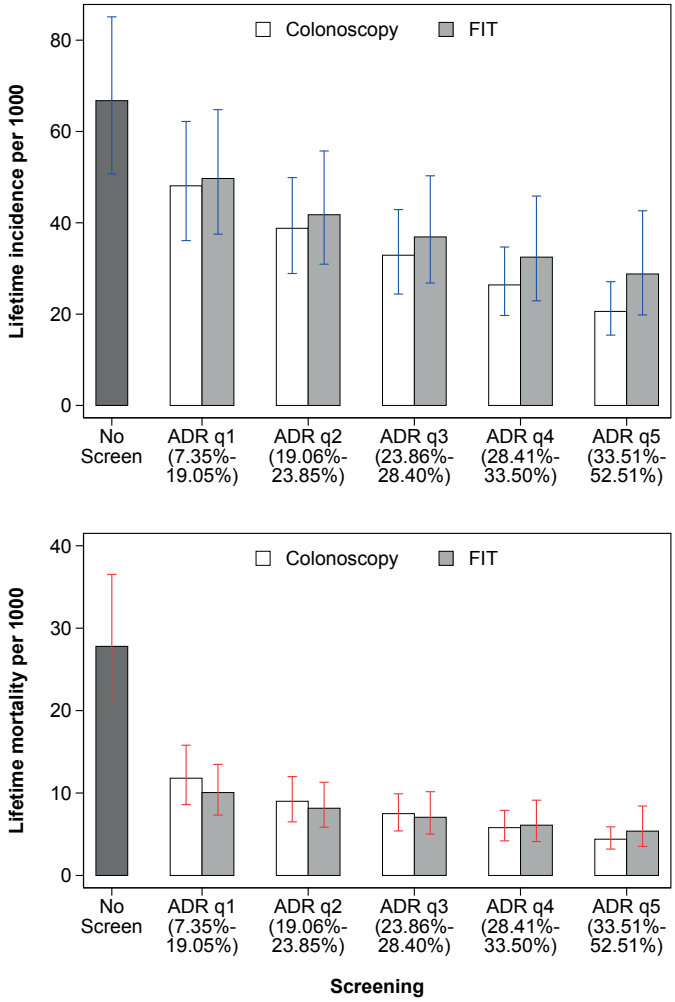


Figure 8.1a-b. Simulated colorectal cancer incidence(a) and mortality(b) per ADR quintile. [a] Abbreviations: ADRqi = Adenoma Detection Rate quintile i (i = 1,…,5). [a] Colonoscopy screening outcomes were previously published.[263] Whiskers represent 95% probability intervals from multivariate probabilistic sensitivity analysis.

The outcomes of FIT screening and primary colonoscopy screening varied according to level of adenoma detection. Among patient receiving FIT screening with potential follow-up colonoscopy from providers in the highest ADR quintile, incidence was 28.8 (95%PI, 19.8-42.6) and mortality 5.4 (95%PI, 3.5-8.4) (**Figure 8.1**). In contrast, for patients receiving colonoscopy screening from the highest ADR quintile providers, the simulated lifetime cancer incidence and mortality were 20.6 (95%PI, 15.4-27.1) and 4.4 (95%PI, 3.2-5.9) per 1000, respectively (relative risks for FIT versus colonoscopy, RR=1.40; 95%PI, 1.09-1.89, and RR=1.22 (95%PI, 0.92-1.75) (**Figure 8.2**). For every 5% point decrease in ADR, simulated incidence was estimated to decrease on average 8.6% (95%PI, 5.5-11.4) for FIT screening and 12.3% (95%PI, 11.1-12.9) for colonoscopy screening (**Table 8.2, Supplementary Figure 8.1**). Thus, in ADR quintile 1, simulated lifetime cancer incidences were more similar, at 49.7 (95%PI, 37.5-64.8) per 1000 for FIT screening and 48.1 (95%PI, 36.1-62.2) per 1000 for colonoscopy screening (RR=1.03; 95%PI, 0.99-1.12) (**Figure 8.1-2**). For every 5% point decrease in ADR, estimated mortality increased by an amount similar to cancer incidence: by 9.4% (95%PI, 6.0-12.7) for FIT screening and 13.3% (95%PI, 11.8-14.2) for colonoscopy screening (**Table 8.2, Supplementary Figure 8.1**). Simulated mortality in quintile 1 was lower with primary FIT than with primary colonoscopy, at 10.1 per 1,000 (95%PI, 7.3-13.5) versus 11.8 per 1000 (95%PI, 8.6-15.8), respectively (RR=0.85; 95%PI, 0.82-0.93).

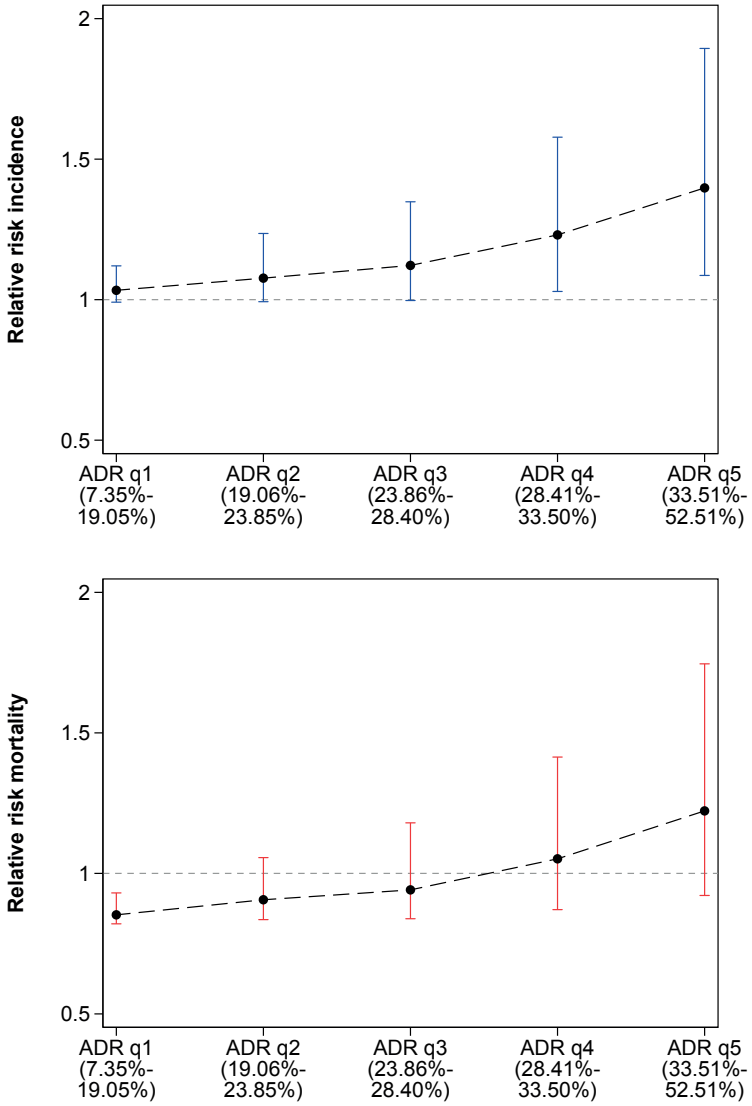


Figure 8.2a-b. Simulated relative risks of colorectal cancer incidence(a) and mortality(b) for FIT versus colonoscopy screening. [a]

Abbreviations: ADRqi = Adenoma Detection Rate quintile i (i = 1,…,5).

[a] Whiskers represent 95% probability intervals from multivariate probabilistic sensitivity analysis. The variable width of confidence intervals from the probabilistic sensitivity analysis was due to interaction of colonoscopy and FIT performance: in the model, lower ADRs decreased the outcome effect of FIT's variable false positive rates and the associated colonoscopy receipt, and higher ADRs increased the effect.

Table 8.2. Sensitivity analysis results: % change in outcomes per 5% point lower ADR. [a]

	Colonoscopy		FIT	
Scenario	Incidence	Mortality	Incidence	Mortality
1. Basecase	12.3	13.3	8.6	9.4
2.a Lower FIT sensitivity [b]			8.0	8.7
2.b Higher FIT sensitivity			9.7	10.8
3.a Lower FIT specificity [c]			10.2	11.3
3.b Higher FIT specificity			7.0	7.6
4.a More emphasized variation in adenoma ≤5mm [d]	9.7	11.3	7.1	7.7
4.b Less emphasized variation in adenoma ≤5mm	12.6	14.1	11.3	12.6

Abbreviations: FIT = Fecal immunochemical test, ADR = adenoma detection rate

[a] Mean simulated outcome differences per 5% decrease in ADR were derived by linear regression and presented relative to the model outcomes for ADR quintile 1 ($5 \times \beta_{ols}/outcome_{q1}$). The actual ADR-outcome relationship was slightly convex (rather than perfectly linear), particularly for FIT screening outcomes: for lower levels of ADR, the outcomes impact of changes in ADRs was somewhat larger (see **Supplementary Figure 8.1**).

[b] We assumed 5% point lower/higher sensitivity of FIT for adenomas, and 10-15% point lower/higher values for cancer.

[c] We assumed 2.5% point lower/higher specificity of FIT.

[d] With more emphasis on small adenomas, all variation in ADRs was attributed to sensitivity of colonoscopy for adenomas smaller than 5 mm, which varied from 5.4%, lowest, to 98%, highest quintile. With less emphasis, all ADR variation was attributed equally to sensitivity for small, medium, and large adenomas, which varied from 26.0% to 98%.

Sensitivity analysis

Outcomes were sensitive to the assumed test characteristics for FIT and colonoscopy (**Table 8.2**). The relative increase in cancer mortality per 5% point lower ADR was smaller for FIT screening when assuming lower FIT sensitivity (8.7%) or higher assumed FIT specificity (7.6%), and larger when assuming higher FIT sensitivity (10.8%) or lower specificity (11.3%). When ADR variation was attributed predominantly to small adenomas, the mortality change was lower for both colonoscopy (11.3%) and FIT (7.7%), while with more variation in detection of larger adenomas variation, mortality changes were large than the base-case (14.1% versus 12.6%, respectively). In all scenarios, the outcome gradient across ADR quintiles was larger for colonoscopy screening than FIT screening.

DISCUSSION

Using microsimulation modeling, we estimated that there is an inverse relationship between physicians' ADR and estimated colorectal cancer screening outcomes that may be stronger when primary screening is performed with colonoscopy than with FIT. Although FIT-based and colonoscopy-based screening strategies were similar on average in terms of their estimated mortality reduction,[259] with providers from the highest ADR quintile, the model suggests that primary colonoscopy screening would result in fewer colorectal cancer cases and deaths than FIT screening. Conversely, FIT screening outperformed colonoscopy in terms of mortality reductions when physician ADRs levels were <20% (male and female patients combined).

The simulated outcome differences between FIT and colonoscopy screening can be explained by the different test characteristics. While colonoscopy, with relatively long screening intervals, provides long-term protection through removal of most existing lesions at the time of screening,[166] the more frequent FIT screening with follow-up colonoscopy of positive results may primarily detect large adenomas and early-stage cancers before they progress to more advanced-stages.[161] The model assumed that physicians with lower detection rates have a higher proclivity for missing small rather than large adenomas.[163,263] Therefore, FIT outcomes were relatively more stable to varying ADRs than primary screening with colonoscopy (9.4% versus 13.3% estimated increase in disease-related mortality per 5% point ADR decrease). In an alternative model with more variation in assumed sensitivity of colonoscopy for large adenomas, outcomes remained more stable for FIT than colonoscopy, but differences were smaller (12.6% versus 14.1% increase in mortality per 5% point ADR decrease).

Another consequence of the different test characteristics of colonoscopy and FIT was that, although FIT was more effective for preventing colorectal cancer deaths than low-quality screening colonoscopy, primary colonoscopy screening resulted in lower estimated colorectal cancer incidence across all ADR quintiles. This is an advantage for colonoscopy screening, which has induced some expert panels to favor colonoscopy over other less invasive modalities for colorectal cancer screening.[115] In contrast, given the different risk and benefit profiles of the different strategies, the most recent recommendation by the United States Preventive Services Task Force (USPSTF) puts more emphasis on patient preferences and shared decision-making.[103]

To our knowledge, the present study is the first to have looked at the influence of ADRs on screening outcomes for a stool-based screening setting. Previous empirical studies have found inverse associations between physician ADR levels and post-colonoscopy cancer risk.[125,127,283] In the largest study to this date, Corley and colleagues found associations between ADR and interval cancer risk that were

similar for screening, diagnostic, and surveillance exams.[124] In a previous modeling study, we estimated that the observed ADR variation may translate to 50-60% differences in lifetime colorectal cancer outcomes for primary colonoscopy screening.[263] The present study suggests that variations in ADR may have less influence on the outcomes of FIT (maximum estimated differences of 42-44%).

The current study may overestimate the differences for fecal testing. While multiple studies have shown that there is substantial variation in ADRs from screening exams,[124,128,275] there are much less data available on variation in adenoma detection during colonoscopies after positive fecal colorectal cancer screening test results. Physicians may examine a patient more carefully with evidence of gastrointestinal blood loss, which may improve the sensitivity of the examination, even for small adenomas that are unlikely to have caused the positive test result. Although adenoma detection rates in diagnostic examinations are not directly comparable to those in screening exams, wide variation in detection rates from population-based FIT screening settings leaves room for substantial variation in miss rates.[288,289] Higher observed risks of cancer after positive FITs followed by negative colonoscopies (for adenomas) could also be indicative of suboptimal quality.

A limitation of this study is the lack of direct experimental data to inform the model on efficacy of FIT and colonoscopy screening.[86,122] We modeled the efficacy of FIT using an established approach used before to inform the United State Preventive Services Task Force.[259] This approach combines evidence from guaiac fecal occult blood testing trial data[190] with observational data on FIT's diagnostic performance.[161] Colonoscopy efficacy estimates were derived similarly using flexible sigmoidoscopy trial results.[117] The simulated mortality effects of FIT are consistent with results from a recent major population-based study,[105] and colonoscopy effects are within the outcome range of observational studies,[107] supporting the use of this approach for the present study.

A strength of this study is that we based our estimates for variable colonoscopy performance characteristics on community-based data regarding interval cancer incidence rates after colonoscopy screening according to physician ADR.[124] Our assumptions have been shown to match well with the observed decreasing incidence pattern from lower to higher ADRs.[124,263] Alternative models with relatively more or less emphasis on variation in detection of diminutive lesions, as evaluated in sensitivity analyses, matched the data less well, which suggests that our base-case assumptions are reasonable. However, we cannot rule out other possible explanations for the observed incidence pattern, such as an association of ADR with adequate polyp management[290] or serrated polyp detection rates.[131]

Our study focused on the influence of observed ADR variation on screening effectiveness. There may be other important, independent, modifiable outcome determi-

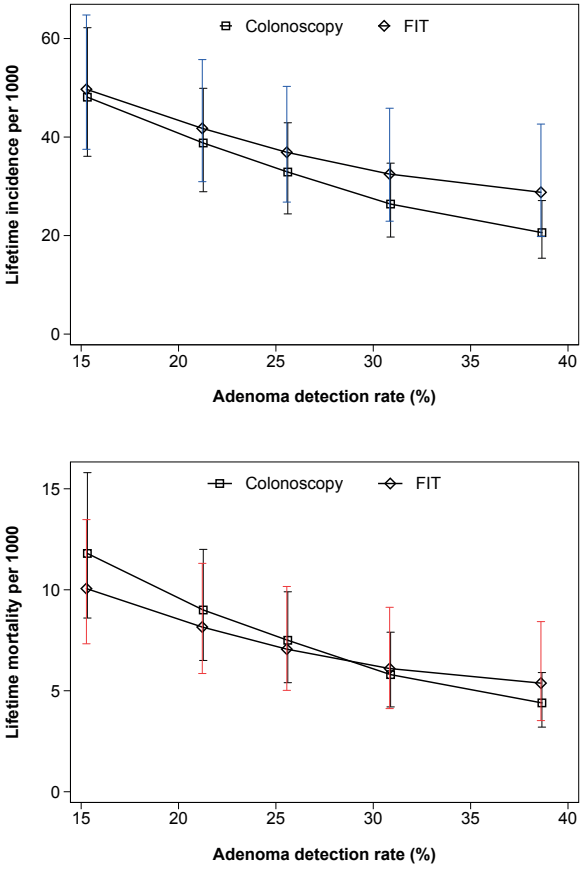
nants for colonoscopy and FIT. Other studies have identified potential determinants in ambient FIT temperature,[291] time from positive FIT to diagnostic colonoscopy follow-up,[292] and particularly, patient adherence.[260] We assumed 100% adherence with both colonoscopy and FIT screening, while in reality adherence may differ for colonoscopy and FIT.[83-86] Patients' willingness to comply with annual fecal colorectal cancer testing methods over longer periods of time is uncertain, with no studies having assessed adherence for >4 subsequent rounds.[90-92] Future modeling studies should assess and rank the relative contribution of all outcome determinants for screening effectiveness to inform priorities of quality-related interventions.

This study has two main implications. First, our results confirm that physician ADR is an important indicator for colorectal cancer screening performance, irrespective of whether the primary screening modality is colonoscopy or FIT. This underscores the importance of ongoing efforts to measure and improve physicians' ADR scores,[131,293] as formalized by some countries in quality assurance programs.[116,294] Recent research suggests that endoscopist training programs may effectively increase ADR levels.[295,296] If large population-based studies confirm that such programs also have a favorable health impact, other screening programs should consider offering similar trainings to stimulate higher ADRs.

Our results further imply that ADR may be useful not only as a quality indicator for screening, but also as a predictor of comparative screening program performance and outcomes. We found that the benefits of FIT relative to colonoscopy screening may differ depending on the quality of colonoscopy achieved in a particular program. In high quality settings, colonoscopy may provide the best possible protection against colorectal cancer deaths, but in settings with lower ADR levels, the more frequently repeated FIT screening may be more effective. This propones colonoscopy quality as one of the relevant factors that policy makers may consider in selecting the most appropriate screening method for their particular setting. Research is needed to assess from what number of exams ADR can be reliably estimated and used as a predictor of both screening outcomes in general, and comparative performance of alternative screening methods in particular.

To conclude, the relative cancer incidence and mortality reductions for FIT versus colonoscopy screening may differ based on colonoscopy quality, as measured by ADR. Although the estimated mortality benefits are similar for FIT and colonoscopy with average ADR levels, colonoscopy screening may result in fewer cancer deaths in settings with higher ADR levels, while FIT screening may result in fewer deaths in lower ADR settings.

APPENDIX 8



Supplementary Figure 8.1a-b Simulated incidence(a) and mortality(b) per ADR level [a]

[a] Colonoscopy screening outcomes were previously published.[263] Whiskers represent 95% probability intervals from multivariate probabilistic sensitivity analysis. Data points on the x-axis represent ADR quintile averages.

Chapter 9

Consequences of time to diagnostic colonoscopy following a positive fecal colorectal cancer screening test

Article published (Almetric = 0):

RG Meester, AG Zauber, CA Doubeni et al.
Clin Gastroenterol Hepatol. 2016 Oct;14(10):1445-1451.e8.

ABSTRACT

BACKGROUND-AIMS: Delays in diagnostic testing after a positive screening test result may undermine the benefits of colorectal cancer (CRC) screening, but little empirical data exist on the harms of such delays. We used microsimulation modeling to evaluate the consequences of time to colonoscopy following a positive fecal immunochemical test (FIT).

METHODS: An established microsimulation model was used to simulate a program of annual FIT screening, with colonoscopy follow-up of positive tests (cutoff 20 μg/g) at various time intervals ≤12 months. Main outcomes were CRC incidence and mortality; additional outcomes were total life-years lost and net costs of screening.

RESULTS: For patients receiving diagnostic colonoscopy within two weeks of a positive FIT, the estimated lifetime CRC incidence and mortality risks were 35.5 and 7.8/1000 patients, respectively. Every added month of time to colonoscopy was associated with 0.1/1000 increased cancer incidence (+0.3 compared to colonoscopy at ≤2 weeks) and mortality (+1.4%). When colonoscopy was received at 12 months after the result date, disease incidence and mortality were 37.0 (+4%) and 9.1/1000 (+16%), respectively. Total years-of-life gained from screening for the entire screening cohort decreased from an estimated 93.7/1000 patients with almost immediate follow-up, to 84.8/1000 (-9%) with follow-up at 12 months, and cost-savings from screening decreased from US $208 to $100 per patient.

CONCLUSION: Modeling suggests that delays of up to 12 months in the follow-up of positive FITs may result, proportionally, in losses of up to nearly 10% in overall screening benefits. This underscores the importance of timely diagnostic follow-up of positive FITs.

INTRODUCTION

As a two-stage screening strategy, the effectiveness of fecal occult blood testing depends on receiving adequate follow-up testing for positive results, generally with colonoscopy. There are no clear guidelines, however, for the appropriate time interval to follow-up colonoscopy. Some studies suggest that intervals of 6 months or longer are common in actual clinical practice.[297,298] United States national patient safety goals emphasize the importance of prompt clinical evaluation of abnormal laboratory test results, but in the case of fecal colorectal cancer testing, the relationship between the time interval from the date of a positive result to diagnostic colonoscopy and colorectal cancer outcomes is not well known. A recent literature review identified two small studies on the subject,[299] the largest of which suggested that longer intervals to receipt of colonoscopy may be associated with higher likelihood of advanced-stage colorectal cancer.[300] However, the study was underpowered to detect statistical differences.

To inform patients, policy, and clinical decision-making on colorectal cancer screening, we used a microsimulation model approach to evaluate the effect of different lengths of time from a positive fecal immunochemical test (FIT) result to receipt of colonoscopy on colorectal cancer incidence, stage distribution, mortality, and the cost-effectiveness of screening programs. In sensitivity analyses we also evaluated other fecal colorectal cancer tests.

METHODS

For this study, we used the MISCAN-colon model (**Chapter 2**) to simulate an average-risk United States population cohort who received annual FIT screening between ages 50-75 years (see Supplementary Table 9.1 for the main model assumptions). For FIT screening, the simulated stage distribution of screen-detected cancers and the simulated mortality effects were consistent with data from population-based studies,[105,301,302] supporting the use of this approach for assessing the effect of lag in diagnostic testing after a positive fecal test result.

Outcomes

Outcomes evaluated were lifetime colorectal cancer incidence, stage and mortality in FIT positive patients for different time intervals to follow-up colonoscopy, as well as the benefits and cost of the FIT screening program as a whole. We also estimated the continuous outcome differences associated with each additional month to colonoscopy using linear regression. Life-years and costs were discounted at the conventional 3% per year.[38]

Analysis

We simulated 10 million men and women born January 1st, 1960. All patients without diagnosed colorectal cancer participated in and complied with annual FIT screening.[114] We considered five scenarios for the average time from positive FIT (OC Sensor, cutoff level for a positive result is 20 μg/g [100 ng/ml]) to follow-up colonoscopy: 2 weeks, 1 month, 2 months, 3 months, 6 months, 12 months and no follow-up colonoscopy. These lag-times were applied to each simulated patient and at every occurrence of a positive result. Patients with adenomas detected at colonoscopy received surveillance colonoscopy per guidelines after 3 – 5 years, depending on the size and multiplicity of adenomas detected.[139]

Sensitivity analysis

In the sensitivity analysis, we evaluated two alternative colorectal cancer screening tests, including the gFOBT (Hemoccult II) and the multi-target stool DNA test (Cologuard) (**Table 9.1**). We further evaluated several alternative model scenarios, including: biennial FIT screening, 50% longer or shorter average duration of the preclinical cancer phase (sojourn time); 5-15 percentage-point lower or higher sensitivity of colonoscopy depending on the lesion size (**Supplementary Table 9.1**) (to account for variation in adenoma detection);[263] 5-15 percentage-point lower or higher sensitivity of FIT (**Supplementary Table 9.1**); 50% lower or higher FIT false-positive rates (1-specificity); and randomly distributed rather than deterministic time to diagnostic follow-up (Gamma[μ,1]; μ = 2 weeks or 3/6/12 months).

Role of the funding source

This study was conducted within the Population-Based Research Optimizing Screening Through Personalized Regimens (PROSPR) consortium, which aims to conduct multisite, coordinated, trans-disciplinary research to evaluate and improve screening and is funded by the NCI. This work is also supported partly by resources from the Veterans Affairs Puget Sound Health Care System.

Results

Colorectal cancer outcomes in FIT positive patients

Among FIT screening participants with a positive test result, the lifetime risks of colorectal cancer incidence and mortality without any diagnostic follow-up were estimated as 82.8 and 34.4 per 1000 patients, respectively. Among patients who had diagnostic colonoscopy within two weeks, the risk of colorectal cancer was reduced to 35.5 per 1000 (**Figure 9.1a**) and the risk of death from colorectal cancer was reduced to 7.8 per 1000 (**Figure 9.1b**). Of the diagnosed cancers, 57% were stage

I, 24% stage II, 12% stage III, and 7% stage IV (**Figure 9.2**). For every additional month to diagnostic colonoscopy, estimated colorectal cancer incidence was higher by 0.1 per 1000 (or a 0.3% relative difference) compared to diagnostic colonoscopy within two weeks, as was cancer-related mortality (1.4% relative difference). For the scenario of diagnostic follow-up at 12 months from a positive FIT, colorectal cancer incidence was 37.0 per 1000, which was about 1.4 cases per 1000 (4%) higher than for almost immediate follow-up, and cancer-related mortality was higher by 1.3 deaths per 1000 (16%). Diagnosed cancers shifted towards more advanced stages, with 50% diagnosed in stage I, 28% stage II, 14% stage III, and 8% stage IV, which is an absolute 7% lower share of stage I cancers.

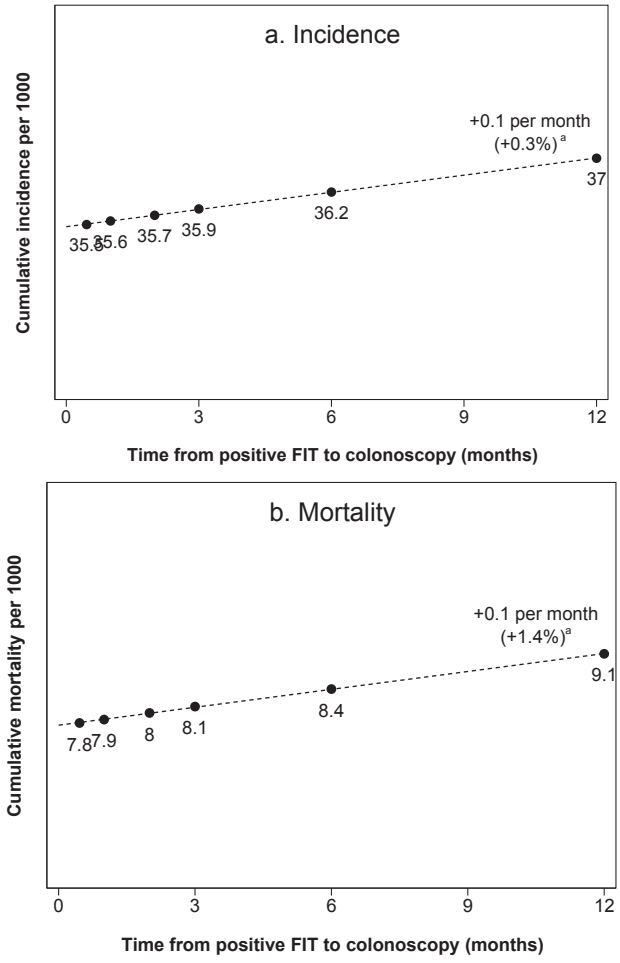


Figure 9.1a,b Lifetime colorectal cancer incidence (a) and mortality (b) in FIT positive patients.
[a] Relative to the scenario of follow-up within two weeks from a positive result.

Total benefits and cost of FIT screening for the entire screening cohort

Among all FIT screening participants, the lifetime risks of colorectal cancer incidence and mortality without any diagnostic follow-up of positive test results - the equivalent of no screening - were 64.8 and 26.8 per 1000, respectively, and 133.5 years of life were lost per 1000 patients due to colorectal cancer (**Table 9.1**). An annual FIT screening program in which diagnostic follow-up of positive tests occurred within two weeks averted 29.2 colorectal cancer cases, 19.4 colorectal cancer deaths and the loss of 93.7 life-years to the disease per 1000 patients. Screening with diagnostic follow-up within 2 weeks of positive results was cost-saving compared to no screening, with a net cost-saving of US $208 per screened patient. With follow-up at 12 months, the number of prevented colorectal cancer cases and deaths decreased to 27.8 and 18.5 per 1000 patients, respectively. Years-of-life saved were 8.9 (9%) lower than with almost immediate follow-up, and at 84.8 per 1000 patients; screening remained cost-saving, but net cost-savings decreased to US $100 per screened patient.

Table 9.1 Simulated cost-effectiveness of FIT screening for the entire screening cohort.

Lifetime outcomes per 1000	Screening						
patients	Average time from positive FIT to colonoscopy (months)						
	None	0 (2 weeks)	1	2	3	6	12
Colorectal cancer outcomes							
Cancer cases	64.8	35.5	35.6	35.7	35.9	36.2	37.0
Advanced cancer cases [a]	53.4	17.1	17.2	17.5	17.7	18.5	19.9
Cancer deaths	26.8	7.4	7.4	7.5	7.6	7.8	8.3
Years of life lost [b]	133.5	39.9	39.9	39.9	42.0	44.4	48.7
Effectiveness of screening							
Cases prevented		29.2	29.2	29.0	28.9	28.5	27.8
Advanced cases prevented [a]		36.3	36.2	35.9	35.7	34.9	33.5
Deaths prevented		19.4	19.4	19.3	19.2	19.0	18.5
Years of life saved [b]		93.7	93.7	93.7	91.5	89.1	84.8
Healthcare costs, US $1000 [b]							
Total costs of screening and treatment	5612	5404	5,411	5,420	5430	5459	5512
Incremental costs to no screening		-208	-201	-193	-182	-153	-100
Cost-effectiveness ratio		c.s.	c.s.	c.s.	c.s.	c.s.	c.s.

Abbreviations: c.s. = cost-saving.

[a] Advanced-stage cancer cases are stage II-IV according to the 5th edition Cancer Staging Manual from the American Joint Committee on Cancer.

[b] Life-years and costs were discounted at the conventional 3% per year.

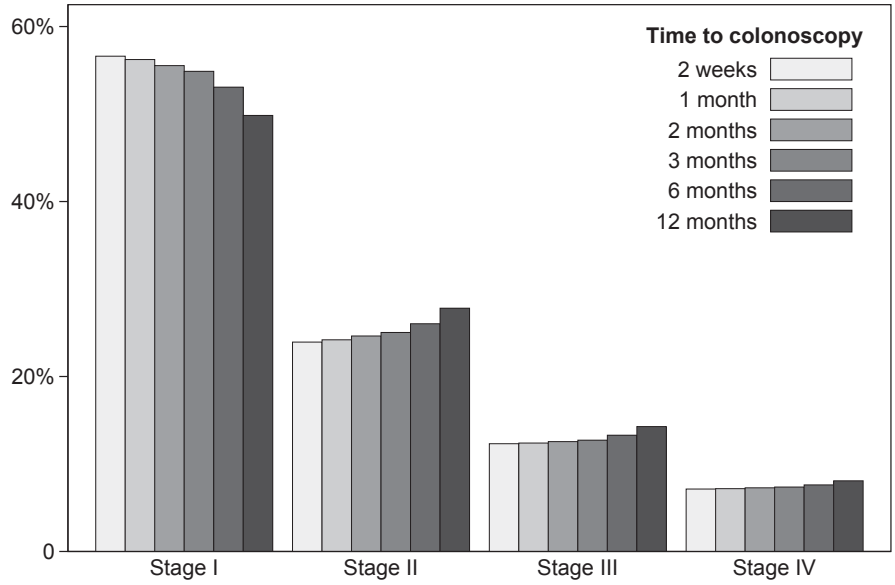


Figure 9.2 Stages of newly diagnosed colorectal cancer cases in FIT positive patients according to time to diagnostic colonoscopy.

Sensitivity analyses

The influence of time to diagnostic testing, per additional month to colonoscopy, was approximately twice as high for gFOBT as for FIT, but was similar for annual FIT versus stool DNA testing every three years (**Figure 9.3**). The results were stable to assumptions on FIT sensitivity and follow-up exam sensitivity for small adenomas, but sensitive to assumptions on test specificity, the length of screening intervals and cancer progression rates. With lower false positive rates or wider screening intervals the effects of time to diagnostic colonoscopy were more than 50% larger than the base-case. A random distribution of time to colonoscopy rather than a deterministic value made hardly any difference for our main outcomes.

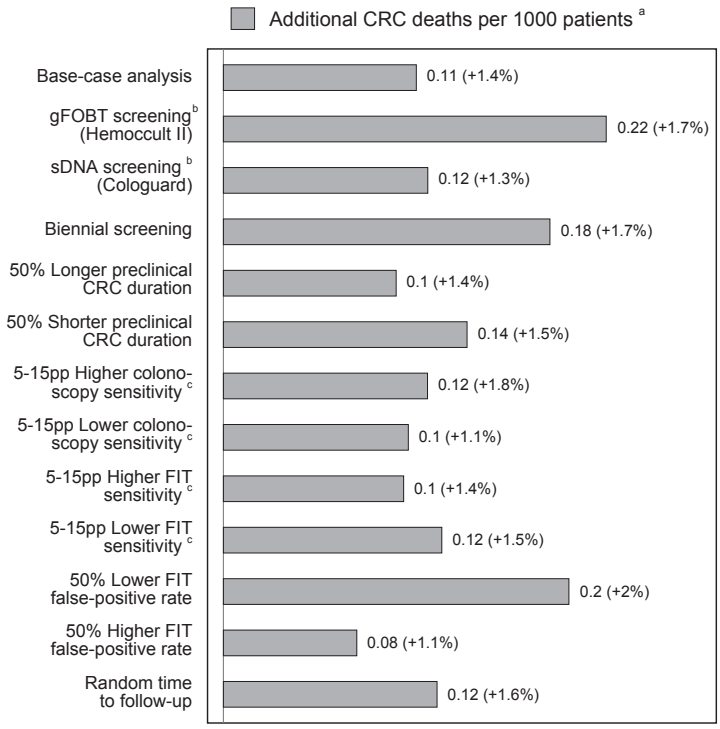


Figure 9.3 Estimated mortality increase per additional month to diagnostic colonoscopy in FIT positive patients, under various scenarios.
Abbreviation: CRC = colorectal cancer; pp = percentage point.
[a] Effects relative to the scenario of follow-up within two weeks from a positive result are presented within parentheses.
[b] See **Supplementary Table 9.1** for the assumed test characteristics.
[c] See **Supplementary Table 9.1** for assumed uncertainty in FIT and colonoscopy sensitivity according to lesion size or stage.

DISCUSSION

In the absence of high-quality observational data, we used an established microsimulation model to estimate the consequences of different times to colonoscopy following a positive FIT for the benefit and cost of colorectal cancer screening. Our results suggest that longer time to follow-up might lead to clinically relevant increases in the risks of colorectal cancer, advanced-stage colorectal cancer and colorectal cancer mortality. Although FIT screening remained cost-saving even with 12 months to follow-up, cancer-related mortality in patients with a positive test

increased more than 15% relatively, and overall life-years gained from screening decreased nearly 10% relatively.

In our analyses, longer time to diagnostic colonoscopy slightly increased the total number of cancers diagnosed due to progression of adenomas to new cancers during that time interval. However, the relative increase in cancer-related mortality was more than three times larger than the relative increase in incidence. This difference stemmed from the relatively slow progression rate of adenomas compared to the more rapid rate of carcinomas progressing from early to more advanced stage disease. Therefore, later follow-up of positive FIT during the year after a positive test influenced the stage of diagnosis more than development of new disease. In our model, the shift to more advanced-stage of diagnosis was the primary driver for the relatively large mortality effect.

Our model results were robust using alternative assumptions regarding the sensitivity of the fecal tests and follow-up exam, but sensitive to assumptions on test specificity, the length of screening intervals and cancer progression rates (**Figure 9.3**). For patients with a positive gFOBT, longer time to follow-up resulted in larger estimated risk increases than the base case with FIT. This mainly reflects the differences in test specificity: using more specific screening tests resulted in a smaller cumulative number of false positive patients without a higher risk of cancer, and consequently, larger, less diluted effects of longer follow-up intervals for the true positive patients. Although we did not evaluate FIT with other-than-standard cutoff levels for positivity (>20 μg/g), by analogy to gFOBT, we would expect larger resulting effects with higher cutoffs, and smaller effects with lower cutoffs. Despite a lower test specificity, stool DNA testing every three years did not result in smaller mortality effects of time to follow-up than annual FIT due to the wider recommended screening intervals.[115] Because of the wider intervals there were fewer total screenings and false positive patients, which offset the effects of lower test specificity. Finally, the duration of the preclinical cancer phase is uncertain,[50,303] and shorter durations increased the likelihood of disease progression and mortality in case of longer time to follow-up.

Some small prior observational studies have estimated the association of time to colonoscopy after positive gFOBT with cancer stage, however, gFOBT has different test characteristics than FIT. One study in 231 subjects found a large, but insignificant relative increase of 7% in the odds of advanced neoplasia (10 mm or more, >25% villous architecture, high-grade dysplasia or intramucosal carcinoma) per additional month to colonoscopy.[300] Although this is larger than the effect we estimated for cancer, the relatively small size of the above study prohibits any meaningful conclusions from such a comparison. Another study in 100 patients found no significant association between time to follow-up and colorectal cancer incidence and mortality.[304] Clearly, both studies were underpowered to detect small-to-moderate effects

on incidence and mortality. In a statistical power analysis we estimated that a case-control design would require at least 3000 cases of advanced-stage colorectal cancer and a history of preceding positive FIT, to demonstrate our model's projected 2.7% estimated relative increase in stage II-IV disease per additional month to colonoscopy, or the corresponding mortality effect (**Appendix 9**, *Statistical Power Analysis*). Other studies have estimated the influence of time to diagnosis for any endoscopy (e.g. for symptoms), and have suggested no, or even inverse, associations,[305-307] but these studies of symptomatic conditions may not be valid for inference of screening tests. In symptomatic patients, disease stage may influence the severity of symptoms, and thereby also the priority for follow-up.

In our analyses, FIT screening was suggested to be highly cost-effective (cost-saving) compared to no screening, similar to other cost-effectiveness studies.[111,266,308] This was mainly due to averted treatment of (advanced-stage) colorectal cancer and the high associated costs. With only one gFOBT trial reporting significant effects on incidence,[156] the effectiveness of FIT for cancer prevention, through the detection and removal of adenomas, is not well established. Superior performance characteristics of FIT to gFOBT-Hemoccult II and less demanding sample collection requirements suggest that FIT could be at least as effective,[114] but no trial data exist.[86] Our approach to estimating FIT efficacy is well-established, and has been used before in the decision analysis to inform the U.S. Preventive Services Task Force.[204] The simulated stage-distribution for screen-detected cancers was consistent with observed data from population-based FIT screening programs,[301,302] as were the estimated mortality effects,[105]

The present study has some limitations. First, in contrast with our assumptions, longer times to colonoscopy may not occur randomly, for example, they may be more common in elderly patients or in patients with comorbid conditions.[298] These patients generally benefit less from screening,[309] and may therefore also have smaller adverse consequences from longer times to examination after a positive FIT. Further, we assumed that false-positive results from asymptomatic benign bleeding occur randomly over individuals and that adenomas are missed randomly, while in reality, false-positive and false-negative results may cluster in specific patients or lesions, e.g. serrated polyps.[56] Because FIT positive patients undergoing a diagnostic examination generally do not return to FIT screening for years, our assumptions may have understated the unknown long-term diagnostic performance of FIT, and therefore the effect of time to diagnostic colonoscopy (**Figure 9.3**).

The findings of this study are applicable primarily to patients who use fecal-based testing methods for colorectal cancer screening. Consequences of time to diagnostic colonoscopy may differ for patients who use a mix of tests for screening, including colonoscopy. Further, we focused our analysis on the effects of time to diagnostic

testing after a positive test result. However, time to therapy in patients with diagnosed cancer may also vary in practice. Thus, future studies are needed to assess the interrelatedness and joint effects of the lag both in diagnostic testing and receipt of treatment on the outcomes of stool-based CRC screening.

To conclude, using modeling we found that deferring diagnostic evaluation may lead to substantial increases in mortality in FIT positive patients. Although the differences between an almost immediate evaluation and an evaluation at up to three months of a positive FIT are small, longer delays in follow-up of up to 12 months may result in more substantial losses, over time, in the overall benefits of screening.

APPENDIX 9

Modeling effects of time to follow-up

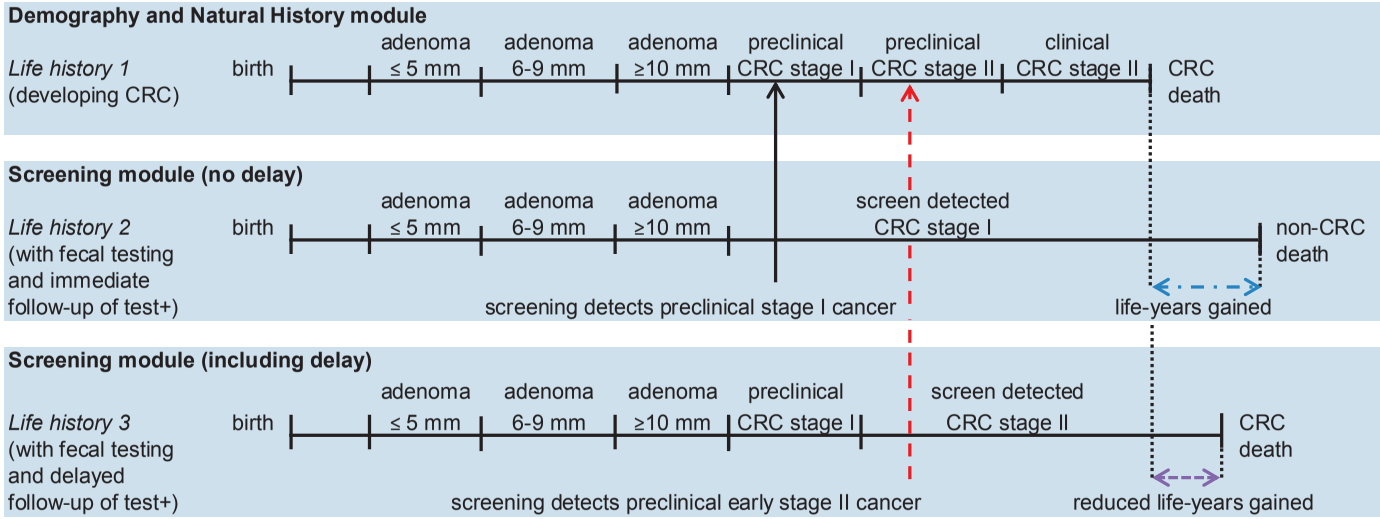
Longer time intervals to diagnostic follow-up of a positive fecal colorectal cancer screening test may lead to more cancers progressing to more advanced-stage, and thereby higher risk of colorectal cancer death. For example, if the simulated Patient A introduced in **Chapter 2** would receive his/her colonoscopy examination at a later point indicated by the red arrow, cancer would progress further to less treatable stage II (*Life history 3,* **Supplementary Figure 9.1**). As a result of this, cancer death would only be deferred, no longer prevented. Hence, for this example patient, the model suggests that longer time to diagnostic follow-up results in a decrease in life-years saved and fewer cancer-related deaths prevented. With random variation in the occurrence and progression of adenoma and cancer progression rates, the effects of time to diagnostic testing vary across different simulated patients.

Statistical power analysis

For this study, we approximated the required case-control study size for statistical demonstration of the simulated effect of time from a positive FIT result to diagnostic colonoscopy on advanced-stage cancer incidence (stage II-IV) and cancer-related mortality, using bootstrap analysis.[310] We assumed 10 years of average patient follow-up time, and a uniform distribution for time intervals of 0 to 12 months. Patients were grouped according to time to follow-up of 0-3 months, 3-6 months, 6-9 and 9-12. The associated risk differences were assessed using generalized linear modeling. The assumed power and significance thresholds were 80% and 5%, respectively.

Statistical demonstration of the relative differences in advanced-stage cancer incidence rates (+2.7% per additional month to colonoscopy in 50-80 year-old patients) was estimated to require 3000 cases with advanced-stage cancer and a preceding positive FIT, and a similar number of random FIT positive patients matched in terms of post-colonoscopy follow-up. Demonstrating the estimated mortality rate difference (+2.3%) also required 3000 cases of deceased colorectal cancer patients with a history of positive FIT, and 3000 controls or random FIT positive patients.

With follow-up intervals of more than 12 months included, the power of case-control studies improved. Assuming follow-up intervals were uniformly distributed from 0 to 24 months, the required number of cases and controls to statistically demonstrate the effect of time to colonoscopy on incidence and mortality would decrease to approximately 1000.



Supplementary Figure 9.1 Integrating MISCAN modules for an example patient.
Abbreviations: CRC = colorectal cancer, FIT = fecal immunochemical test.

Supplementary Table 9.1 Test performance assumptions in MISCAN.[a]

Performance characteristic, % [range evaluated in sensitivity analyses]	Colorectal cancer screening test			
	Colono-scopy	FIT (OC Sensor, cutoff >20 µg/g)	gFOBT (HemoccultII)	sDNA (Cologuard)
Sensitivity per lesion				
Adenomas ≤ 5 mm [b]	75 [+/-15]	0.0	0.0	0.0
Adenomas 6 - 9 mm [b]	85 [+/-10]	4.9 [+/-5]	1.3	22.0
Adenomas ≥ 10 mm [b]	95 [+/-5]	16.2 [+/-5]	6.5	28.4
Stage I-IV cancer long before the occurrence of clinical symptoms[c]	95 [+/-5]	64[+/-15]	18.2	86.4
Stage I-IV cancer shortly before the occurrence of clinical symptoms[c]	95 [+/-5]	89[+/-10]	50.8	96.7
Specificity [d]	100	95[+/-2.5]	97.5	90
Completeness colonoscopy [e]	98	-	-	-
Complication rate colonoscopy				
with polypectomy, age 50-100y	0.4-8.5	-	-	-
fatal complications	0.0033	-	-	-
without polypectomy [f]	-	-	-	-

Abbreviations: FIT = fecal immunochemical test; gFOBT = guaiac fecal occult blood test; sDNA = multi-target stool DNA test.

[a] For references see **Supplementary Table 9.2**.

[b] The adenoma sensitivity estimates and uncertainty range for FIT (and sDNA) were obtained by calibrating our model outcomes to the estimated per-person sensitivities from Imperiale et al.[161] The per-person sensitivity of FIT for adenomas, advanced adenomas, and cancer was 7.6, 23.8, 73.8, respectively. Colonoscopy sensitivity estimates were derived from a systematic review,[163] and uncertainty was assumed to be larger for small adenomas to reflect adenoma detection rate variation.[263]

[c] We assumed that fecal occult blood testing is more sensitive in cancers towards the end of the occult invasive period (close, time-wise, to becoming symptomatic): for preclinical cancers which will become symptomatic within the same stage, assumed test sensitivity was higher. This assumption showed good concordance with guaiac fecal occult blood test trial results.[190]

[d] The probability of a false positive result was random in the base-case analysis, and independent of person or lesion. We assumed perfect specificity for colonoscopy including pathological examination of detected lesions. We included costs for pathological examination of non-adenomatous lesions (e.g. hyperplastic -) in 15% of exams not detecting adenomas.

[e] This is the proportion of colonoscopies visualizing the maximum point of reach of the endoscope, i.e. the cecum.

[f] Colonoscopy without polypectomy was not associated with a higher risk of complications. The risk of complications for polypectomy increased exponentially with age. Complications include serious GI events such perforation and gastrointestinal bleeding requiring blood transfusions; other GI events such as paralytic ileus, nausea, vomiting and dehydration, abdominal pain; and cardiovascular events such as myocardial infarction or angina, arrhythmias, congestive heart failure, cardiac or respiratory arrest, or syncope, hypotension, or shock. The fatal perforation rate was derived from estimates of the incidence of perforation and case-fatality for perforation.[271 272]

Supplementary Table 9.2 Key assumptions in MISCAN-Colon.

Parameter	Value	Reference
Demography		
All-cause mortality	United States lifetables	CDC 2010
Natural history		
Adenoma onset	Nonhomogeneous Poisson process: Exponential(μλ) time to event, μ=260-18y for age 25-80y, μ=18-170y for age 80-100y, λ=Gamma(1;2) risk factor	[181,182,185] et al.
Adenoma progression		[117]
State transitions	0-89% adenomas progressive for age 0-100y 30% size 6-9mm progress to cancer 70% size 6-9mm first become 10+mm	
State duration, y (total)	Exponential(130)	
Preclinical cancer progression		[157,188,189]
Stage transitions [a]	0-31% stage I become clinical for age 0-100y 18-58% stage II become clinical for age 0-100y 58-49% stage III become clinical for age 0-100y	
Stage durations, y (average)	Exponential(2.5)	
Colorectal cancer incidence (without exposure to screening)	See Figure 2.3-4	SEER 1975-1979
5y Colorectal cancer survival [b]	58-71% stage I, depending on location 58-62% stage II, depending on location 33% stage III 6% stage IV	SEER 2000-2010
Screen test performance		
Sensitivity FIT, colonoscopy	See Supplementary Table 9.1	[161,163,262,311]
Specificity FIT, colonoscopy		[161,262,269,270,311]
Completeness colonoscopy		[286,287]
Complication rate colonoscopy		[271,272]
Costs, US $ *		
FIT	26	[152]
Colonoscopy		
without polypectomy	890	
with polypectomy	1099	
Complications	6051	[277]
Per life-year with cancer care, stage I-IV †		[278]
Initial	34,116-69,978	
Continuing	2718-10,177	
Terminal, cancer death	58,531-80,555	
Terminal, other-cause death	18,063-45,939	

Supplementary Table 9.2 Legend Key assumptions in MISCAN-Colon.
Abbreviations: FIT = Fecal Immunochemical Test; CDC = United States Centers for Disease Control and Prevention; SEER = Surveillance Epidemiology and End Results program.
[a] There is additional variation in cancer stage transitions according to bowel location.
[b] These cancer survival rates are adjusted for competing risks.
[c] Screen- and treatment costs include patient time costs (opportunity costs of spending time on screening or being treated for a complication or colorectal cancer), but not cost of traveling, lost productivity, and unrelated health care or non-health care costs in added years-of-life. Patient time was valued at May 2014 median Unites States wage ($17.09 per hour), and assumed to be zero for FIT, 8 hours for colonoscopy, and 16 hours for complications. Cost estimates for life-years with cancer care by Yabroff et al. already included patient time.[278]
[d] Care for colorectal cancer was divided in three clinically relevant phases: the initial, continuing, and terminal care phase. 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 cancer patients dying from the disease and cancer 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.

PART IV

Cost-Effectiveness of Personalized Regimens

Chapter 10

Should colorectal cancer screening be considered in elderly persons without previous screening? A cost-effectiveness analysis

Article published (AltMetric not assessed):

F van Hees, JD Habbema, RG Meester et al.
Ann Intern Med. 2014 Jun 3;160(11):750-9.

ABSTRACT

BACKGROUND: The U.S. Preventive Services Task Force recommends against routine screening for colorectal cancer (CRC) in adequately screened persons older than 75 years but does not address the appropriateness of screening in elderly persons without previous screening.

OBJECTIVE: To determine at what ages CRC screening should be considered in unscreened elderly persons and to determine which test is indicated at each age.

DESIGN: Microsimulation modeling study.

DATA SOURCES: Observational and experimental studies.

TARGET POPULATION – TIME HORIZON – PERSPECTIVE: Unscreened persons aged 76 to 90 years with no, moderate, and severe comorbid conditions. Lifetime horizon. Societal perspective.

INTERVENTION: One-time colonoscopy, sigmoidoscopy, or fecal immunochemical test (FIT) screening.

Outcome Measures: Quality-adjusted life-years gained, costs, and costs per quality-adjusted life-year gained.

RESULTS OF BASE-CASE ANALYSIS: In unscreened elderly persons with no comorbid conditions, CRC screening was cost-effective up to age 86 years. Screening with colonoscopy was indicated up to age 83 years, sigmoidoscopy was indicated at age 84 years, and FIT was indicated at ages 85 and 86 years. In unscreened persons with moderate comorbid conditions, screening was cost-effective up to age 83 years (colonoscopy indicated up to age 80 years, sigmoidoscopy at age 81 years, and FIT at ages 82 and 83 years). In unscreened persons with severe comorbid conditions, screening was cost-effective up to age 80 years (colonoscopy indicated up to age 77 years, sigmoidoscopy at age 78 years, and FIT at ages 79 and 80 years).

RESULTS OF SENSITIVITY ANALYSIS: Results were most sensitive to assuming a lower willingness to pay per quality-adjusted life-year gained.

LIMITATION: Only persons at average risk for CRC were considered.

CONCLUSION: In unscreened elderly persons CRC screening should be considered well beyond age 75. A colonoscopy is indicated at most ages.

INTRODUCTION

In its most recent recommendation statement on colorectal cancer (CRC) screening, the U.S. Preventive Services Task Force (USPSTF) recommends screening using fecal occult blood testing, sigmoidoscopy, or colonoscopy, starting at age 50 years and continuing up to age 75 years.[246] The USPSTF recommends against routine screening in persons older than 75 years with an adequate screening history.[246] This latter recommendation is warranted by an analysis showing that the benefits of continuing screening from age 50 to 85 years instead of 75 years do not justify the additional colonoscopies required.[204] Although the USPSTF did not address the appropriateness of screening in inadequately screened elderly persons, this recommendation has led many members of the medical community to believe that no one older than 75 years should be screened for CRC.[312,313] However, because unscreened elderly persons are at greater risk for CRC than adequately screened elderly persons, screening them is likely to be effective and cost-effective up to a more advanced age. If so, the lack of more specific recommendations on the age to stop screening may result in an unfounded denial of access to screening in elderly persons who were never screened for CRC—a group representing 23% of all U.S. persons older than 75 years.[89]

Many other elderly persons continue to be screened up to their late 80s or early 90s.[63] However, at these ages, screening is not likely to be cost-effective, even in those without previous screening. First, the high risk for death of competing disease at advanced age tends to offset the benefits of screening.[314,315] Second, the risks for screening-induced harms (colonoscopy-related complications and overdiagnosis and overtreatment of CRC) increase with increasing age.[271]

The objective of this study was to determine up to what age CRC screening should be considered in elderly persons without previous screening and to determine which screening test—a colonoscopy, sigmoidoscopy, or fecal immunochemical test (FIT)—is indicated at what age. We performed separate analyses for elderly persons with no, moderate, and severe comorbid conditions because the effectiveness and cost-effectiveness of screening depend heavily on a person's life expectancy.

METHODS

We used Microsimulation Screening Analysis–Colon (MISCAN-Colon) to quantify the effectiveness and costs of screening (Chapter 2).

Populations Simulated

For each age between 76 and 90 years, we simulated a cohort of 10 million elderly persons without previous screening with no, moderate, and severe comorbid conditions (a total of 45 cohorts). Compared with cohorts of adequately screened elderly persons, the risk for CRC in these cohorts was substantially greater: CRC and adenomas were prevalent in 0.3% and 14.1%, respectively, of simulated persons aged 80 years with negative screening colonoscopies at ages 50, 60, and 70 years; these lesions were prevalent in 2.6% and 44.9%, respectively, of simulated patients aged 80 years without previous screening.

We used comorbid condition level–specific life tables to simulate elderly persons with no, moderate, and severe comorbid conditions (27). Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction; as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS; and as having no comorbid conditions if none of these conditions are present.

Screening Strategies

We simulated 1-time colonoscopy, 1-time sigmoidoscopy, and 1-time FIT screening within each cohort. Test characteristics and complication rates for each screening test are given in **Supplementary Table 10.1**. Patients with an adenoma or CRC detected during sigmoidoscopy or with a positive FIT result were referred for a diagnostic colonoscopy. Persons with adenomas detected and removed during a screening or diagnostic colonoscopy were assumed to have colonoscopy surveillance according to the current guidelines.[139] We assumed that surveillance continued until the diagnosis of CRC or death. Adherence to screening and diagnostic and surveillance colonoscopies was assumed to be 100%.

We restricted ourselves to 1-time colonoscopy and 1-time sigmoidoscopy screening because performing more screening colonoscopies or sigmoidoscopies is unlikely to be cost-effective at older age. We explored the effect of FIT screening during 2 consecutive years in a sensitivity analysis.

Utility Losses Associated With CRC Screening

We assumed a utility loss (that is, a loss of quality of life) equal to 2 full days of life per colonoscopy (0.0055 quality-adjusted life-years [QALYs]), 1 day of life per sigmoidoscopy (0.0027 QALYs), and 2 weeks of life per complication (0.0384 QALYs) (**Table 10.1**).[278,316,317] We also assigned a utility loss to each life-year (LY) with CRC care.[316]

Table 10.1 Utility Losses and Costs Associated With CRC Screening

Variable	Initial Care	Continuing Care	Terminal Care, *CRC Death*	Terminal Care, *Other-Cause Death*
Utility loss, *QALY*[a]				
Per FIT	0			
Per sigmoidoscopy				
Without biopsy	0.0027			
With biopsy	0.0027			
Per colonoscopy				
Without polypectomy/biopsy	0.0055			
With polypectomy/biopsy	0.0055			
Per complication of colonoscopy	0.038			
Per LY with CRC care [b c]				
Stage I CRC	0.12	0.05	0.70	0.05
Stage II CRC	0.18	0.05	0.70	0.05
Stage III CRC	0.24	0.24	0.70	0.24
Stage IV CRC	0.70	0.70	0.70	0.70
Costs, [d]				
Per FIT	42			
Per sigmoidoscopy				
Without biopsy	299			
With biopsy	557			
Per colonoscopy				
Without polypectomy/biopsy	887			
With polypectomy/biopsy	1096			
Per complication of colonoscopy	6045			
Per LY with CRC care [b]				
Stage I CRC	36 683	3050	63 809	19 176
Stage II CRC	49 234	2870	63 555	17 279
Stage III CRC	59 759	4021	67 041	21 457
Stage IV CRC	77 790	12 178	88 368	49 866

← **Table 10.1 Legend** Utility Losses and Costs Associated With CRC Screening
Abbreviations: CRC = colorectal cancer; FIT = fecal immunochemical test; LY = life-year; QALY = quality-adjusted life-year.
[a] The loss of quality of life associated with a particular event.
[b] Care for CRC was divided in 3 clinically relevant phases: the initial, continuing, and terminal care phase. The initial care phase was defined as the first 12 mo after diagnosis; the terminal care phase was defined as the final 12 mo of life; the continuing care phase was defined as all months in between. In the terminal care phase, we distinguished between patients with CRC who were dying of CRC and those dying of another cause. For patients surviving less than 24 mo, the final 12 mo were allocated to the terminal care phase and the remaining months were allocated to the initial care phase.
[c] Utility losses for LYs with initial care were derived from a study by Ness et al.[316] For LYs with continuing care for stage I and II CRC, we assumed a utility loss of 0.05 QALYs; for LYs with continuing care for stage III and IV CRC, we assumed the corresponding utility losses for LYs with initial care. For LYs with terminal care for CRC, we assumed the utility loss for LYs with initial care for stage IV CRC. For LYs with terminal care for another cause, we assumed the corresponding utility losses for LYs with continuing care.
[d] Costs are presented in 2013 U.S. dollars and include copayments and patient time costs (i.e., the opportunity costs of spending time on screening or being treated for a complication or CRC) but do 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: $16.71/h.[317] We assumed that FITs, sigmoidoscopies, colonoscopies, and complications used up 1, 4, 8, and 16 h of patient time, respectively. Patient time costs were already included in the estimates for the costs of LYs with CRC care obtained from a study by Yabroff et al.[278]

The assignment of utility losses to LYs with CRC care works 2 ways: On the 1 hand, screening prevents cancer by the detection and removal of adenomas, thereby reducing LYs with CRC care and hence resulting in a gain of quality of life. On the other hand, screening results in overdiagnosis and overtreatment of cancer, resulting in LYs with CRC care in persons who would never have been diagnosed with CRC without screening and hence a loss of quality of life. The net effect on quality of life depends on the balance between cancer cases prevented and cancer cases overdiagnosed and can be either positive or negative.

Costs Associated With CRC Screening

The cost-effectiveness analyses were conducted from a societal perspective. The costs of colonoscopy, sigmoidoscopy, and FIT were based on 2007 Medicare payment rates and copayments (**Table 10.1**).[152,317] The costs of complications were obtained from a cost analysis of cases of unexpected hospital use after endoscopy in 2007.[277] We added patient time costs to both. The costs of LYs with CRC care were obtained from an analysis of SEER–Medicare linked data and included copayments and patient time costs.[278] We adjusted all costs to reflect the 2013 level using the U.S. consumer price index.[318] The assignment of costs to LYs with CRC care also works 2 ways: On the 1 hand, screening prevents cancer, reducing the costs of CRC care. On

the other hand, screening results in overtreatment of cancer, increasing these costs. The net effect can be either a reduction or an increase in costs.

Outcomes

For each cohort, we quantified the effectiveness (that is, the number of CRC cases prevented, CRC deaths prevented, LYs gained, and QALYs gained) and costs of 1-time colonoscopy, sigmoidoscopy, and FIT screening, applying the conventional 3% annual discount rate for both.

Analyses

We first determined the cost-effectiveness of each screening strategy compared with no screening for all cohorts. For each comorbidity level, we determined the upper age at which each screening strategy was cost-effective compared with no screening, assuming a willingness to pay per QALY gained of $100 000.

We subsequently performed an analysis to determine the optimal screening strategy for each cohort (that is, the most effective, still cost-effective screening strategy). To do so, we first excluded all dominated screening strategies (that is, those that were more costly and less effective than combinations of other strategies). We determined the incremental cost-effectiveness ratio for all remaining strategies (the so-called efficient strategies): the additional costs per additional QALY gained compared with the next-less-effective and costly strategy. From the efficient strategies, we selected the optimal strategy, again assuming a willingness to pay per QALY gained of $100 000.

Sensitivity Analyses

We repeated our analyses, assuming one half and twice the base-case utility losses for colonoscopy, sigmoidoscopy, and complications; a utility loss of 0.12, 0.18, 0.24, and 0.70 QALYs for each LY with continuing care for stage I, II, III, and IV CRC, respectively; 25% higher and 25% lower costs for colonoscopy, sigmoidoscopy, and FIT; 25% higher and 25% lower costs for CRC care; twice the base-case miss rates for adenomas and CRC for both sigmoidoscopy and colonoscopy; no surveillance in patients with adenomas; a 25% higher and a 25% lower risk for CRC in all cohorts; and a willingness to pay per QALY gained of $50 000. Further, we explored the effect of FIT screening during 2 consecutive years.

This study did not include patient-specific information and was exempt from institutional review board review.

Role of the Funding Source

The study was supported by the National Cancer Institute. The funding source had no role in the study's design, conduct, and reporting.

RESULTS

Effectiveness

The effectiveness of CRC screening in unscreened elderly persons declined with increasing age (**Table 10.2**, **Supplementary Table 10.2**). For example, 1-time colonoscopy screening prevented fewer CRC deaths (4.5 vs. 11.9 per 1000 persons) and resulted in fewer LYs gained (12.3 vs. 68.5 per 1000 persons) in healthy persons aged 90 years than in healthy persons aged 76 years. Moreover, whereas colonoscopy screening prevented 15.4 CRC cases per 1000 persons aged 76 years, it resulted in overdiagnosis and hence overtreatment of 7.7 CRC cases per 1000 persons aged 90 years. As a result, colonoscopy screening resulted in a positive overall effect on length and quality of life (that is, a net health benefit) in healthy persons aged 76 years (67.2 QALYs gained per 1000 persons) but in a net harm in healthy persons aged 90 years (1.7 QALYs lost per 1000 persons).

One-time sigmoidoscopy and, particularly, 1-time FIT screening were generally less effective than 1-time colonoscopy screening (**Table 10.2**): For example, in healthy persons aged 76 years, colonoscopy screening resulted in 67.2 QALYs gained per 1000 persons, whereas sigmoidoscopy and FIT screening resulted in 53.9 and 24.2 QALYs gained per 1000 persons, respectively. The only exceptions were seen at the most advanced ages, at which FIT screening was most effective—a result primarily explained by the 0 utility loss associated with this test. In persons with moderate and, particularly, severe comorbid conditions, screening was less effective than in persons without comorbid conditions (**Supplementary Table 10.3**).

Table 10.2 The Effectiveness of 1-Time Colonoscopy, Sigmoidoscopy, and FIT Screening in Elderly Persons Without Previous Screening With No Comorbid Conditions [a]

Screening Strategy,	CRC cases Prevented, *n*[b]	CRC Deaths Prevented, *n*	LYs Gained, *n*[c]	Effect on Quality of Life, *QALY*[d]					QALYs Gained, *n*[f]
by Age				Screening Tests	Diagnostic Exams	Surveillance Exams	Compli-cations	LYs With [e] CRC Care	
1-time COL									
76y [g]	15.4	11.9	68.5	−5.5	0	−3.2	−0.6	8.1	67.2
80y	10.4	10.7	52.9	−5.5	0	−2.8	−0.7	3.0	46.9
85y	0.8	7.4	28.3	−5.5	0	−2.0	−0.9	−2.9	17.1
90y	−7.7	4.5	12.3	−5.5	0	−1.4	−1.1	−6.1	−1.7
1-time FSIG									
76y	12.0	9.4	54.6	−2.7	−1.6	−2.2	−0.4	6.2	53.9
80y	8.2	8.7	43.1	−2.7	−1.7	−2.0	−0.5	2.3	38.6
85y	0.6	6.0	23.1	−2.7	−1.7	−1.4	−0.6	−2.3	14.3
90y	−6.2	3.7	9.9	−2.7	−1.6	−1.0	−0.7	−4.9	−1.0
1-time FIT									
76y	1.7	4.1	25.9	0	−0.4	−0.5	−0.1	−0.6	24.2
80y	0.2	4.2	22.5	0	−0.4	−0.4	−0.1	−2.2	19.2
85y	−2.8	3.4	13.8	0	−0.5	−0.4	−0.1	−3.8	9.0
90y	−6.2	2.3	6.6	0	−0.5	−0.3	−0.2	−4.7	0.9

Abbreviations: COL = colonoscopy; FSIG = flexible sigmoidoscopy; FIT = fecal immunochemical test; CRC = colorectal cancer; LY = life-year; QALY = quality-adjusted life-year.

[a] Results are based on comparison with no screening, with results per 1000 persons and discounted by 3%. Persons are classified as having no comorbid conditions if none of the following conditions are present: an ulcer, a history of acute myocardial infarction, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, cerebrovascular disease, chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS.

[b] Negative values occur when the number of CRC cases prevented by screening is exceeded by the number of CRC cases overdiagnosed by screening.

[c] The effect of screening on quantity of life.

[d] The effect of the screening test, diagnostic colonoscopies, surveillance colonoscopies, complications, and LYs with CRC care on quality of life. Values are derived by multiplying number(s) of events with the corresponding utility loss(es) per event stated in **Table 10.1**. For example, when applying the 1-time colonoscopy screening strategy, 1000 persons have a screening colonoscopy in each cohort. Because the utility loss per screening colonoscopy is 0.0055 QALYs, the total utility loss due to screening colonoscopies is 5.5 QALYs in each cohort.

[e] Screening results in a gain of quality of life by preventing LYs with CRC care and a loss of quality of life by adding LYs with CRC care. The net effect can be a gain of quality of life (positive values) or a loss of quality of life (negative values). As a result of the shift from preventing to overdiagnosing CRC with increasing age, the net effect on quality of life becomes less favorable with age. Whereas 1-time colonoscopy screening in unscreened elderly without comorbid conditions reduced the total number of LYs with CRC care for stage III or IV CRC at age 76 y (−14 LYs per 1000 persons), it increased this number of LYs at age 90 y (+16 LYs per 1000 persons).

[f] Calculated by adding LYs gained and QALYs for screening tests, diagnostic colonoscopies, surveillance colonoscopies, complications, and LYs with CRC care. The effect of screening on quantity and quality of life incorporated in 1 measure (i.e., the net health benefit of screening). Discrepancies between the columns may occur due to rounding.

[g] More detailed results for this cohort are given in **Supplementary Table 10.2**.

Costs

Whereas the effectiveness of screening in unscreened elderly persons declined with increasing age, the net costs of screening increased substantially (**Table 10.3**). Although colonoscopy screening was associated with a lifetime cost of $725 000 per 1000 healthy persons aged 76 years, it was associated with a lifetime cost of $2 130 000 per 1000 healthy persons aged 90 years. This increase was again explained by the shift from preventing to overtreating CRC with age.

Table 10.3 Costs of 1-Time Colonoscopy, Sigmoidoscopy, and FIT Screening in Elderly Persons Without Previous Screening With No Comorbid Conditions [a]

Screening Strategy	Cost (Thousands), US$					
	Screening Tests [b]	Diagnostic Exams	Surveillance Exams	Complications	LYs With CRC Care [c]	Total [d]
1-time COL						
76y‖	983	0	569	98	−925	725
80y	987	0	484	114	−483	1102
85y	987	0	350	137	230	1705
90y	986	0	239	168	737	2130
1-time FSIG						
76y‖	387	309	397	64	−718	439
80y	392	331	345	75	−380	764
85y	392	330	251	89	189	1251
90y	390	323	169	106	592	1580
1-time FIT						
76y [e]	42	80	88	14	−7	218
80y	42	87	78	17	130	355
85y	42	93	62	23	356	577
90y	42	98	46	29	541	756

Abbreviations: CRC = colorectal cancer; FIT = fecal immunochemical test.

[a] Results are based on comparison with no screening, with results per 1000 persons and discounted by 3%. Persons are classified as having no comorbid conditions if none of the following conditions are present: an ulcer, a history of acute myocardial infarction, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, cerebrovascular disease, chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS.

[b] At very advanced age, the costs of screening colonoscopies and sigmoidoscopies show a slight decline. This is explained by the small observed decrease in the prevalence of adenomas at very advanced age (11–18, 20, 21).

[c] Screening prevents costs by preventing LYs with CRC care and induces costs by adding LYs with CRC care. The net effect can be an increase in costs (positive values) or a decrease in costs (negative values).

[d] Discrepancies between the columns may occur due to rounding.

[e] More detailed results for this cohort are given in **Supplementary Table 10.2**.

Besides being the most effective strategy, colonoscopy screening was also the most expensive (**Table 10.3**). For example, in healthy persons aged 76 years, the costs of colonoscopy screening were $725 000 per 1000 persons compared with $439 000 and $218 000 for sigmoidoscopy and FIT screening, respectively. In persons with moderate and, particularly, severe comorbid conditions, screening was not only less effective but also more costly (**Supplementary Table 10.4**).

Cost-Effectiveness Compared With No Screening

As the effectiveness of screening declined with increasing age and the costs increased substantially, the cost-effectiveness of screening deteriorated rapidly with age (**Figure 10.1**). In unscreened elderly persons without comorbid conditions, colonoscopy and sigmoidoscopy screening were cost-effective up to age 85 years, whereas FIT screening was cost-effective up to age 86 years. In elderly persons with moderate comorbid conditions, colonoscopy and sigmoidoscopy screening were cost-effective up to age 82 years, whereas FIT screening was cost-effective up to age 83 years. In persons with severe comorbid conditions, colonoscopy and sigmoidoscopy screening were cost-effective up to age 79 years, whereas FIT screening was cost-effective up to age 80 years.

Incremental Cost-Effectiveness

We determined the optimal screening strategy for each cohort on the basis of the incremental cost-effectiveness ratios of the efficient screening strategies. In unscreened elderly persons with no comorbid conditions, colonoscopy screening was most effective and still cost-effective up to age 83 years (**Supplementary Table 10.5**, and **Figure 10.2**), sigmoidoscopy screening was the optimal strategy at age 84 years, and FIT screening was the optimal strategy at ages 85 and 86 years. In elderly persons with moderate comorbid conditions, colonoscopy screening was the optimal strategy up to age 80 years, sigmoidoscopy screening was the optimal strategy at age 81 years, and FIT screening was the optimal strategy at ages 82 and 83 years. In persons with severe comorbid conditions, colonoscopy screening was the optimal strategy up to age 77 years, followed by sigmoidoscopy screening at age 78 years and FIT screening at ages 79 and 80 years.

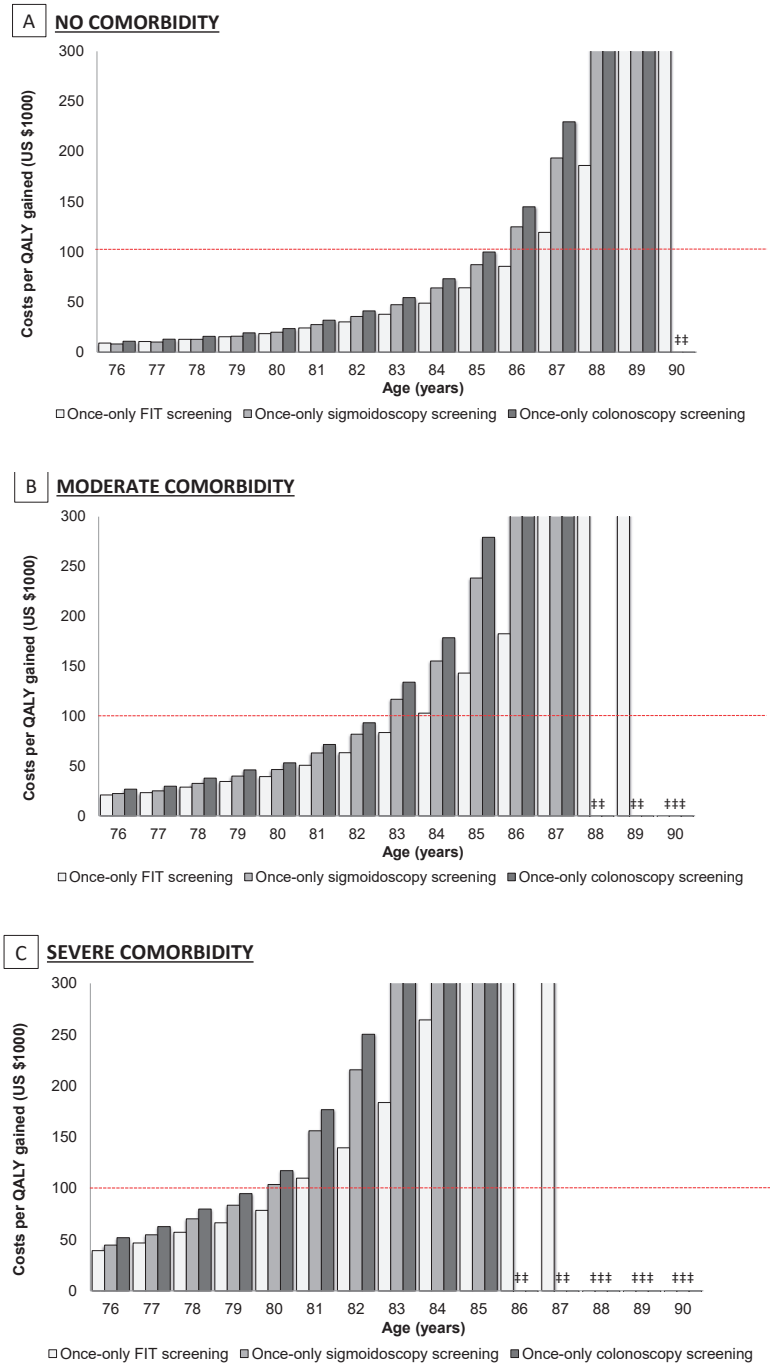


Figure 10.1 a-c The Cost-Effectiveness of Once-Only Colonoscopy, Sigmoidoscopy, and FIT Screening Compared with No Screening in Elderly Without Prior Screening with No (A), Moderate (B), and Severe Comorbidity (C) (3% discounted).[a b]

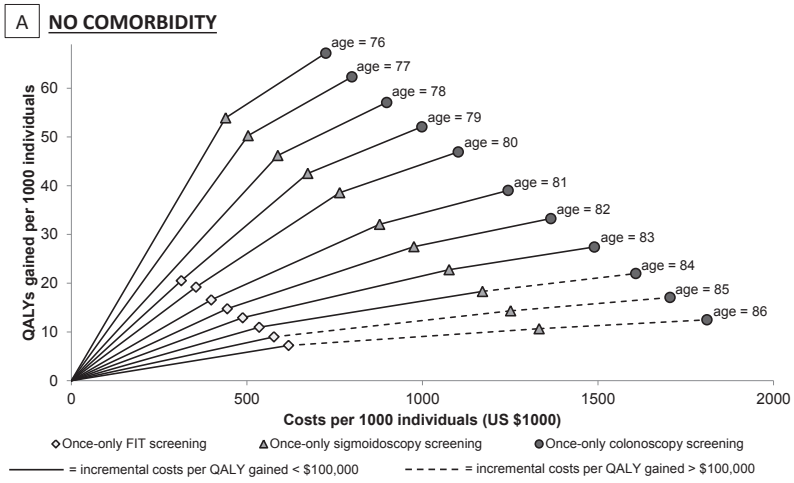
← **Figure 10.1 Legend**

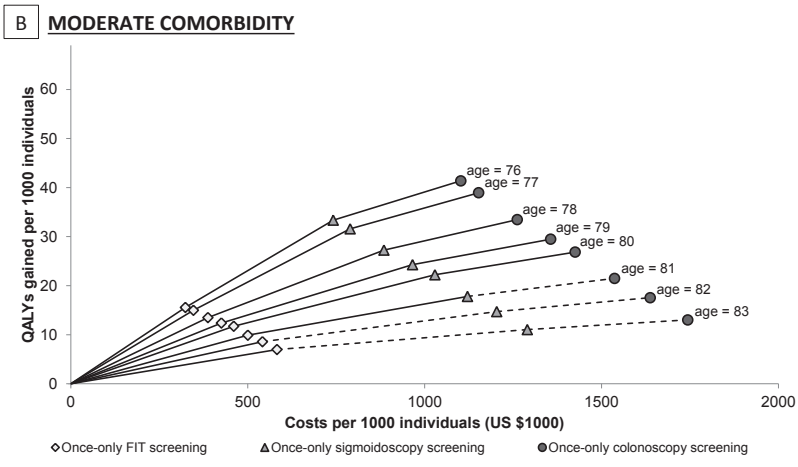
Abbreviations: QALY = quality-adjusted life-year; FIT = fecal immunochemical test.

[a] Individuals are classified as having moderate comorbidity if diagnosed with an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction; as having severe comorbidity if diagnosed with constructive obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS; and as having no comorbidity if none of these conditions is present.

[b] The dashed red line indicates a threshold for the willingness to pay per QALY gained of $100,000. Screening strategies costing less than $100,000 per QALY gained are considered cost-effective.

[c] ‡‡ Signs indicate ages at which screening is associated with a net health loss, rather than a benefit (**Table 10.2**, **Supplementary Table 10.3**).





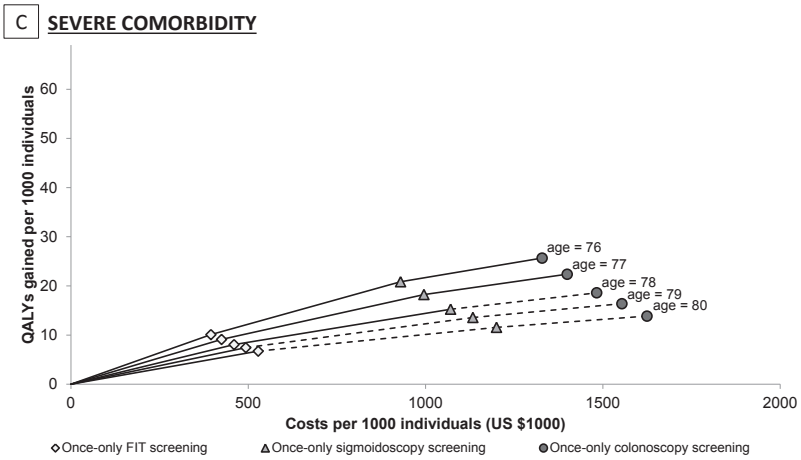


Figure 10.2 a-c The Incremental Costs-Effectiveness of the Efficient Screening Strategies in Elderly Without Prior Screening with No (A), Moderate (B), and Severe Comorbidity (C) (results per 1,000 individuals; 3% discounted).[a b c]

Abbreviations: QALY = quality-adjusted life-year; FIT = fecal immunochemical test

[a] Individuals are classified as having moderate comorbidity if diagnosed with an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction; as having severe comorbidity if diagnosed with constructive obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS; and as having no comorbidity if none of these conditions is present.

[b] In elderly without prior screening with no, moderate, and severe comorbidity, none of the screening strategies is cost-effective from age 87, 84, and 81 onwards, respectively (**Figure 10.1**).

[c] For each age, the efficient screening strategies are connected by an efficiency frontier. A dashed line indicates that the incremental cost-effectiveness ratio of a screening strategy exceeds $100,000 per QALY gained, implying that the strategy is no longer considered cost-effective.

Sensitivity Analyses

Besides comorbid condition level, the upper age at which screening was cost-effective was most sensitive to lowering the threshold for the willingness to pay per QALY gained to $50 000 (**Supplementary Table 10.6**). Based on this threshold, screening unscreened elderly persons with no, moderate, and severe comorbid conditions should be considered up to age 84, 80, and 77 years, respectively. The upper ages at which screening should be considered were robust to all other sensitivity analyses (**Supplementary Table 10.6**).

The tests that were indicated at specific ages differed substantially between analyses (**Supplementary Table 10.6**). Besides the threshold for the willingness to pay per QALY gained, the level of CRC risk and the utility losses associated with colonoscopy, sigmoidoscopy, and complications were the most important factors in this respect.

In persons aged 84 years without comorbid conditions and persons aged 78 years with severe comorbid conditions, sigmoidoscopy screening was not cost-effective compared with FIT screening during 2 consecutive years. In persons aged 85 years without comorbid conditions, persons aged 82 years with moderate comorbid conditions, and persons aged 79 and 80 years with severe comorbid conditions, FIT screening during 2 consecutive years was cost-effective compared with 1-time FIT screening.

DISCUSSION

Our study shows that in elderly persons without previous CRC screening, screening remains cost-effective well beyond age 75 years, which is the recommended age to discontinue screening in adequately screened persons (**Table 10.4**). In unscreened elderly persons with no comorbid conditions, screening was cost-effective up to age 86 years. Screening with colonoscopy was most effective and still cost-effective up to 83 years, sigmoidoscopy was indicated at age 84 years, and FIT was indicated at ages 85 and 86 years. In unscreened elderly persons with moderate comorbid conditions, screening was cost-effective up to age 83 years (colonoscopy indicated up to age 80 years, sigmoidoscopy at age 81 years, and FIT at ages 82 and 83 years). In persons with severe comorbid conditions, screening was cost-effective up to age 80 years (colonoscopy indicated up to age 77 years, sigmoidoscopy at age 78 years, and FIT at ages 79 and 80 years).

In the special situation when an elderly person is willing to have only one type of screening test, the cost-effectiveness of that test compared with no screening becomes relevant. In such a person without comorbid conditions, colonoscopy and sigmoidoscopy screening can be considered up to age 85 years and FIT screening can be considered up to age 86 years. The ages for similar persons with moderate comorbid conditions are 82 years for colonoscopy and sigmoidoscopy and 83 years for FIT; for persons with severe comorbid conditions, the ages are 79 years for colonoscopy and sigmoidoscopy and 80 years for FIT.

Although the incidence of CRC increases up to very advanced ages,[319] the effectiveness of screening declines with increasing age. This decline is primarily explained by the increasing risk for other-cause death. with age, which reduces both the probability that screening will prevent CRC death and the number of LYs gained if death

is prevented. Moreover, the risks for screening-induced harms (colonoscopy-related complications and, more importantly, overdiagnosis and overtreatment of CRC) increase with age.[271] At the same time, the shift from preventing to overtreating CRC causes the net costs of screening to increase with age. Together, these phenomena explain the rapid deterioration of the cost-effectiveness of screening with increasing age.

Table 10.4 Results Summary of CRC Screening Indicated in Elderly Persons Without Previous Screening

Comorbid Condition Status [a]	Age Up to Which CRC Screening Should Be Considered, *y*	Screening Strategy Indicated, by Age										
		76y	77y	78y	79y	80y	81y	82y	83y	84y	85y	86y
No comorbid conditions	86	COL	COL	COL	COL	COL	COL	COL	COL	FSIG	FIT	FIT
Moderate comorbid conditions	83	COL	COL	COL	COL	COL	FSIG	FIT	FIT			
Severe comorbid conditions	80	COL	COL	FSIG	FIT	FIT						

Abbreviations: COL = 1-time colonoscopy; CRC = colorectal cancer; FSIG = 1-time sigmoidoscopy; FIT = 1-time fecal immunochemical test.

[a] Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction; as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS; and as having no comorbid conditions if none of these conditions are present.

Although colonoscopy every 10 years, sigmoidoscopy every 5 years, and FIT every year are almost equally effective when applied from age 50 to 75 years (1, 2), colonoscopy screening is more effective than sigmoidoscopy and FIT when only one screening examination is performed because of its greater overall sensitivity for adenomas and CRC. However, because colonoscopy is also more expensive than sigmoidoscopy and FIT and because the effectiveness of all screening tests is marginal at very advanced ages, screening with colonoscopy is not cost-effective compared with sigmoidoscopy and FIT at the most advanced ages at which screening should be considered.

Screening remains cost-effective up to a more advanced age in persons without comorbid conditions than in those with comorbid conditions because their more favorable life expectancy increases the probability that screening will prevent CRC, thus increasing the effectiveness of screening while simultaneously reducing the costs of CRC care.

To our knowledge, our study is the first to investigate the net health benefit and the cost-effectiveness of CRC screening in persons older than 75 years without previous screening. An earlier study by Ko and Sonnenberg[314] demonstrated that the effectiveness of screening for preventing CRC death declines with increasing age, whereas the probability of screening-related complications increases with age. Further, a study by Lin and colleagues[315] demonstrated that the number of LYs gained by screening declines with age, resulting in an increase in the number of colonoscopies required per LY gained. However, neither study considered costs or measured the overdiagnosis and overtreatment of cancer, which is the most important adverse effect of screening in elderly persons. As a result, these studies cannot easily be used to determine whether unscreened elderly persons should be screened. Some other, more recent studies have suggested that screening should be continued after age 75 years.[312,313] However, these studies did not distinguish between adequately screened elderly persons and elderly persons without previous screening. Further, these studies based their conclusions only on CRC incidence data.

The USPSTF selected its recommended screening strategies based on the number of colonoscopies required per LY gained (undiscounted),[204,246] but we based our conclusions on the costs per QALY gained (discounted at 3% per year). We did so for 2 reasons. First, policymakers should be able to compare the efficiency of a wide range of health interventions; the USPSTF outcome measure does not allow for this. Second, we believe that effects on both length and quality of life should be considered. However, the two approaches led to screening recommendations associated with similar numbers of colonoscopies per LY gained: Screening with colonoscopy as recommended by the USPSTF (that is, at ages 50, 60, and 70 years) required 30 to 35 colonoscopies per LY gained.[204] Also, screening with colonoscopy in unscreened persons aged 83 years with no comorbid conditions, for example, required 32 colonoscopies per LY gained.

Our study has two main limitations. First, we did not perform separate analyses by sex and race. However, we do not expect that results from such analyses would have differed much from the results presented in this paper because a substantial part of the difference in life expectancy between men and women and between blacks and whites is explained by differences in the prevalence of moderate and severe comorbid conditions. Also, persons with the most favorable life expectancy (that is, white females) are at lowest risk for CRC and vice versa. Hence, the effect of life expectancy on the cost-effectiveness of screening is counterbalanced by the effect of CRC risk (at least partially).[320] Second, we did not perform separate analyses for identifiable high-risk subgroups, such as elderly persons with a family history of CRC.[321] In some of these subgroups, screening may be cost-effective up to a more advanced age.

Our analysis highlights some future research directions. First, future research should determine the optimal number of FIT screenings in elderly persons who are relatively young and not willing to have a screening colonoscopy or sigmoidoscopy. Second, other research should study how the benefits, burden, and harms of screening affect patient decisions about CRC screening. Third, studies evaluating the appropriate age to stop screening by comorbid condition level are also required for adequately screened persons.

In conclusion, our study demonstrates that in the 23% of U.S. elderly persons without previous screening, CRC screening should be considered well beyond age 75 years. In unscreened elderly persons with no comorbid conditions, CRC screening should be considered up to age 86 years (up to age 83 years for those with moderate comorbid conditions and up to age 80 years for those with severe comorbid conditions). Screening with colonoscopy is indicated at most ages.

APPENDIX 10

Supplementary Table 10.1 Test Characteristics of Colonoscopy, Sigmoidoscopy, and FIT

Test Characteristic	Test		
	Colonoscopy	Sigmoidoscopy	FIT
Specificity, %	90 [a]	92 [a]	98 [b]
Sensitivity, %			
Small adenomas (≤5 mm)	75 [c]	75 [c]	0 [b]
Medium-sized adenomas (6–9 mm)	85 [c]	85 [c]	5 [b]
Large adenomas (≥10 mm)	95 [c]	95 [c]	26 [b]
CRCs that would not have been clinically detected in their current stage	95 [c]	95 [c]	41 [b]
CRCs that would have been clinically detected in their current stage	95 [c]	95 [c]	77 [b]
Reach	95% cecum	6% splenic flexure [d] 88% sigmoid-descending flexure	-
Complication rate			
Positive result	Increases exponentially with age (0.002-0.048 for 76-90 y) [e]	0	0
Negative result	0	0	0
Mortality rate			
Positive result	0.033/1000 [f]	0	0
Negative result	0	0	0

Abbreviations: CRC = colorectal cancer; FIT = fecal immunochemical test.

[a] We assumed that in 10% of all negative colonoscopy results and in 8% of all negative sigmoidoscopy results a non-adenomatous lesion was detected, resulting in a polypectomy or a biopsy, respectively.

[b] The sensitivity of colonoscopy and sigmoidoscopy for the detection of adenomas and CRC within the reach of the endoscope was obtained from a systematic review on miss rates seen in tandem colonoscopy studies.[163]

[c] The test characteristics of FIT were fitted to the positivity rates and detection rates seen in the first screening round of the Dutch screening trial. We assumed that the probability that a CRC bleeds and thus the sensitivity of FIT for CRC depends on the time until clinical diagnosis, in concordance with our findings for guaiac fecal occult blood test.[190]

[d] The reach of sigmoidoscopy was obtained from a study by Painter et al.[322]

[e] Age-specific risks for complications of colonoscopy requiring a hospital admission or emergency department visit were obtained from a study by Warren et al.[271]

[f] The mortality rate associated with colonoscopies with a polypectomy was derived by multiplying the risk for a perforation obtained from a study by Warren et al[271] by the risk for death given a perforation obtained from a study by Gatto et al.[272]

Supplementary Table 10.2 Effects of 1-Time Colonoscopy Screening in Persons Aged 76 Years Without Previous Screening With No Comorbid Conditions [a]

Effect	Screening	No Screening	Screening − No Screening [b]
Effects on health care use, *n*			
Colonoscopies			
Screening with polypectomy/biopsy	461	0	461
Screening without polypectomy/biopsy	539	0	539
Surveillance with polypectomy/biopsy	219	0	219
Surveillance without polypectomy/biopsy	370	0	370
Complications of colonoscopy	16.2	0	16.2
LYs with initial CRC care [c]			
Stage I	11.5	6.4	5.1
Stage II	8.0	12.4	−4.4
Stage III	5.1	7.3	−2.2
Stage IV	0.7	2.9	−2.2
LYs with continuing CRC care			
Stage I	92.8	34.9	57.9
Stage II	60.0	61.6	−1.6
Stage III	33.9	30.7	3.2 [d]
Stage IV	1.5	5.2	−3.7
LYs with terminal care, ending in CRC death			
Stage I	0.5	0.7	−0.2
Stage II	1.0	2.6	−1.6
Stage III	1.5	3.2	−1.8
Stage IV	1.1	5.8	−4.7
LYs with terminal care, ending in other-cause death			
Stage I	8.3	5.1	3.2
Stage II	5.4	9.3	−4.1
Stage III	2.9	4.6	−1.7
Stage IV	0.2	1.0	−0.8
Effects on health			
CRC cases, *n*	27.9	43.4	−15.5
CRC deaths, *n*	4.5	16.4	−11.9
LYs lost due to CRC, *n*	32.5	100.9	−68.4 [e]
Utility losses, *QALYs*			
Screening colonoscopies	5.5	0	5.5
Surveillance colonoscopies	3.2	0	3.2
Complications of colonoscopy	0.6	0	0.6
LYs with CRC care	25.7	33.8	−8.1
Total	35.1	33.8	1.3
QALYs lost (LYs lost due to CRC + total utility loss), *n*	67.5	134.7	−67.2 [f]

Supplementary Table 10.2 Effects of 1-Time Colonoscopy Screening in Persons Aged 76 Years Without Previous Screening With No Comorbid Conditions [a] (Continued)

Effect	Screening	No Screening	Screening − No Screening [b]
Effects on costs (thousands), *$*			
Screening colonoscopy	983	0	983
Surveillance colonoscopy	569	0	569
Complications of colonoscopy	98	0	98
LYs with CRC care	2404	3329	−925
Total	4054	3329	725 [g]

Abbreviations: CRC = colorectal cancer; LY = life-year; QALY = quality-adjusted life-year.

[a] Results per 1000 persons, discounted at 3% per year. Persons are classified as having no comorbid conditions if none of the following conditions are present: an ulcer, a history of acute myocardial infarction, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, cerebrovascular disease, constructive obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS.

[b] Discrepancies between columns may occur due to rounding.

[c] Because screening results in prevention and earlier detection of CRC, it reduces the total numbers of LYs with initial care for CRC, terminal care for CRC, and terminal care for other causes in patients with CRC; however, because screening improves the average survival of patients with CRC, it increases the total number of LYs with continuing care for CRC.

[d] The increase in LYs with continuing care for stage III CRC is explained by the more favorable average survival that we model for screen-detected vs. clinically detected cancer as described in **Chapter 2**.

[e] The number of LYs gained by screening **(Table 10.2)**.

[f] The number of QALYs gained by screening **(Table 10.2)**.

[g] The costs of screening **(Table 10.1).**

Supplementary Table 10.3 Effectiveness of 1-Time Colonoscopy, Sigmoidoscopy, and FIT Screening in Elderly Persons Without Previous Screening With Moderate and Severe Comorbid Conditions [a]

Screening Strategy, by Comorbidity Level	CRC Cases Prevented, *n*	CRC Deaths Prevented, *n*	LYs Gained, *n*[b]	Effect on Quality of Life, *QALYs*[c]					QALYs Gained, *n*[d]
				Screening Tests	Diagnostic Exams	Surveillance Exams	Complications	LYs With CRC Care [e]	
Moderate comorbidity									
1-time COL									
76 y		9.0	46.3	−5.5	0	−2.6	−0.6	+3.8	41.4
80 y		8.1	35.2	−5.5	0	−2.2	−0.7	0.0	26.8
85 y		5.6	18.9	−5.5	0	−1.6	−0.8	−4.2	6.8
90 y		3.5	8.8	−5.5	0	−1.1	−1.0	−6.1	−4.8
1-time FSIG									
76 y		7.2	36.9	−2.7	−1.6	−1.8	−0.4	+2.9	33.4
80 y		6.6	28.7	−2.7	−1.7	−1.6	−0.4	−0.1	22.2
85 y		4.6	15.4	−2.7	−1.7	−1.1	−0.5	−3.4	5.9
90 y		2.9	7.1	−2.7	−1.6	−0.7	−0.6	−4.9	−3.5
1-time FIT									
76 y		3.3	17.9	0	−0.4	−0.4	−0.1	−1.5	15.6
80 y		3.4	15.4	0	−0.4	−0.4	−0.1	−2.8	11.7
85 y		2.7	9.4	0	−0.5	−0.3	−0.1	−4.0	4.6
90 y		1.9	4.8	0	−0.5	−0.2	−0.2	−4.4	−0.5
Severe comorbidity									
1-time COL									
76 y	2.6	6.7	32.3	−5.5	0	−2.0	−0.5	+1.4	25.7
80 y	−2.2	5.9	23.3	−5.5	0	−1.6	−0.6	−1.7	13.9
85 y	−9.4	4.0	12.2	−5.5	0	−1.1	−0.8	−4.5	0.4
90 y	−14.6	2.6	5.8	−5.5	0	−0.7	−1.0	−5.7	−7.1

Supplementary Table 10.3 Effectiveness of 1-Time Colonoscopy, Sigmoidoscopy, and FIT Screening in Elderly Persons Without Previous Screening With Moderate and Severe Comorbid Conditions [a] (continued)

Screening Strategy, by Comorbidity Level	CRC Cases Prevented, *n*	CRC Deaths Prevented, *n*	LYs Gained, *n*[b]	Effect on Quality of Life, *QALYs*[c]					QALYs Gained, *n*[d]
				Screening Tests	Diagnostic Exams	Surveillance Exams	Complications	LYs With CRC Care [e]	
1-time FSIG									
76 y	2.0	5.3	25.8	−2.7	−1.6	−1.4	−0.3	+1.1	20.8
80 y	−1.9	4.8	19.0	−2.7	−1.7	−1.2	−0.4	−1.4	11.6
85 y	−7.6	3.3	10.0	−2.7	−1.7	−0.8	−0.5	−3.6	0.6
90 y	−11.7	2.1	4.6	−2.7	−1.6	−0.5	−0.6	−4.5	−5.4
1-time FIT									
76 y	−2.2	2.5	12.7	0	−0.4	−0.3	−0.1	−1.8	10.1
80 y	−4.2	2.5	10.4	0	−0.4	−0.3	−0.1	−2.9	6.7
85 y	−7.1	2.0	6.2	0	−0.5	−0.2	−0.1	−3.7	1.7
90 y	−9.5	1.4	3.2	0	−0.5	−0.1	−0.2	−4.0	−1.7

Abbreviations: CRC = colorectal cancer; COL = colonoscopy; FSIG = flexible sigmoidoscopy; FIT = fecal immunochemical test; LY = life-year; QALY = quality-adjusted life-year.

[a] Results are based on a comparison with no screening, given per 1000 persons, and discounted by 3% per year. Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction and as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS. Negative values occur when the number of CRC cases prevented by screening is exceeded by the number of CRC cases overdiagnosed by screening.

[b] The effect of screening on quantity of life.

[c] The effect of the screening test, diagnostic colonoscopies, surveillance colonoscopies, complications, and LYs with CRC care on quality of life. Values are derived by multiplying number(s) of events with the corresponding utility loss(es) per event stated in **Table 10.1**. An example: When applying the 1-time colonoscopy screening strategy, 1000 persons have a screening colonoscopy in each cohort. Because the utility loss per screening colonoscopy is 0.0055 QALYs, the total utility loss due to screening colonoscopies is 5.5 QALYs in each cohort.

[d] The effect of screening on quantity and quality of life incorporated in 1 measure (i.e., the net health benefit of screening), calculated by adding LYs gained and all effects on quality of life. Discrepancies between the columns may occur due to rounding.

[e] Screening results in a gain of quality of life by preventing LYs with CRC care and a loss of quality of life by adding LYs with CRC care. The net effect can be a gain of quality of life (positive values) or a loss of quality of life (negative values). As a result of the shift from preventing to overdiagnosing CRC with increasing age, the net effect on quality of life becomes less favorable with age.

Supplementary Table 10.4 Costs of 1-Time Colonoscopy, Sigmoidoscopy, and FIT Screening in Elderly Persons Without Previous Screening With Moderate and Severe Comorbid Conditions [a]

Screening Strategy, by Comorbidity Level	Cost (Thousands), $					
	Screening Tests [b]	Diagnostic Exams	Surveillance Exams	Compli-cations	LYs With CRC Care [c]	Total [d]
Moderate comorb.						
1-time COL						
76 y	983	0	462	90	−434	1102
80 y	987	0	388	106	−57	1425
85 y	987	0	278	131	502	1898
90 y	986	0	185	161	838	2170
1-time FSIG						
76 y	387	309	323	58	−336	742
80 y	392	331	278	69	−41	1029
85 y	392	330	199	84	409	1414
90 y	390	323	132	100	673	1618
1-time FIT						
76 y	42	80	72	13	116	324
80 y	42	87	63	16	252	460
85 y	42	93	50	22	448	655
90 y	42	98	36	28	578	782
Severe comorb.						
1-time COL						
76 y	983	0	354	83	−91	1329
80 y	987	0	288	99	250	1625
85 y	987	0	199	123	658	1967
90 y	986	0	131	154	868	2139
1-time FSIG						
76 y	387	309	248	52	−67	930
80 y	392	331	206	63	207	1200
85 y	392	330	143	77	534	1477
90 y	390	323	94	95	698	1600
1-time FIT						
76 y	42	80	56	12	204	395
80 y	42	87	47	15	337	528
85 y	42	93	36	20	493	685
90 y	42	98	26	27	576	769

Abbreviations: CRC = colorectal cancer; COL = colonoscopy; FSIG = flexible sigmoidoscopy; FIT = fecal immunochemical test; LY = life-year, comorb. = comorbidity.

[a] Results are based on a comparison with no screening, given per 1000 persons, and discounted by 3% per year. Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction and as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS.

[b] At very advanced age, the costs of screening colonoscopies and sigmoidoscopies show a slight decline. This is explained by the small observed decrease in the prevalence of adenomas at very advanced age.

[c] Screening prevents costs by preventing LYs with CRC care and induces costs by adding LYs with CRC care. The net effect can be an increase in costs (positive values) or a decrease in costs (negative values).

[d] Discrepancies between the columns may occur due to rounding.

Supplementary Table 10.5 ICERs of the Efficient Screening Strategies in Elderly Persons Without Previous Screening by Comorbid Condition Level [a]

Screening Strategy, by Comorbidity Level and Age [b]	QALYs Gained. *N*[c]	Incremental QALYs Gained, [d]	Cost (Thousands), *$*[c]	Incremental Cost (Thousands), *$*[d]	ICER (Thousands), *$/QALY* [e]	Optimal Screening Strateg[f]
No comorbidity						
76 y[g]						
FSIG	53.9	53.9	439	439	8	
COL	67.2	13.3	725	285	21	X
77 y[g]						
FSIG	50.3	50.3	503	503	10	
COL	62.3	12.1	799	296	24	X
78 y[g]						
FSIG	46.2	46.2	588	588	13	
COL	57.1	10.9	898	310	28	X
79 y						
FIT	20.5	20.5	313	313	15	
FSIG	42.5	22.0	673	360	16	
COL	52.1	9.6	998	325	34	X
80 y						
FIT	19.2	19.2	355	355	18	
FSIG	38.6	19.4	764	409	21	
COL	46.9	8.4	1102	338	40	X
81 y						
FIT	16.6	16.6	398	398	24	
FSIG	32.1	15.5	878	480	31	
COL	39.0	7.0	1244	366	52	X
82 y						
FIT	14.8	14.8	444	444	30	
FSIG	27.5	12.7	976	532	42	
COL	33.3	5.8	1365	390	67	X
83 y						
FIT	12.9	12.9	488	488	38	
FSIG	22.8	9.9	1076	588	59	
COL	27.4	4.7	1490	414	88	X
84 y						
FIT	11.0	11.0	535	535	49	
FSIG	18.3	7.3	1171	636	87	X
COL	22.0	3.7	1608	437	118	
85 y						
FIT	9.0	9.0	577	577	64	X
FSIG	14.3	5.3	1251	674	127	
COL	17.1	2.7	1705	454	168	

Supplementary Table 10.5 ICERs of the Efficient Screening Strategies in Elderly Persons Without Previous Screening by Comorbid Condition Level [a] (continued)

Screening Strategy, by Comorbidity Level and Age [b]	QALYs Gained. *N*[c]	Incremental QALYs Gained, [d]	Cost (Thousands), *$*[c]	Incremental Cost (Thousands), *$*[d]	ICER (Thousands), *$/QALY* [e]	Optimal Screening Strateg[f]
86 y						
FIT	7.2	7.2	619	619	86	X
FSIG	10.7	3.4	1332	714	210	
COL	12.5	1.8	1810	478	266	
Moderate comorbidity						
76 y						
FIT	15.6	15.6	324	324	21	
FSIG	33.4	17.8	742	418	23	
COL	41.4	8.0	1102	361	45	X
77 y						
FIT	15.0	15.0	347	347	23	
FSIG	31.6	16.6	789	443	27	
COL	38.9	7.4	1153	363	49	X
78 y						
FIT	13.5	13.5	387	387	29	
FSIG	27.3	13.8	885	497	36	
COL	33.5	6.2	1262	377	61	X
79 y						
FIT	12.4	12.4	426	426	34	
FSIG	24.3	11.9	966	540	45	
COL	29.5	5.2	1356	390	75	X
80 y						
FIT	11.7	11.7	460	460	39	
FSIG	22.2	10.5	1029	569	54	
COL	26.8	4.6	1425	396	86	X
81 y						
FIT	9.9	9.9	500	500	51	
FSIG	17.8	7.9	1121	621	79	X
COL	21.5	3.7	1537	416	112	
82 y						
FIT	8.6	8.6	542	542	63	X
FSIG	14.7	6.1	1204	662	109	
COL	17.6	2.8	1638	434	155	
83 y						
FIT	7.0	7.0	583	583	83	X
FSIG	11.0	4.1	1290	707	172	
COL	13.0	2.0	1744	453	227	

Supplementary Table 10.5 ICERs of the Efficient Screening Strategies in Elderly Persons Without Previous Screening by Comorbid Condition Level [a] (continued)

Screening Strategy, by Comorbidity Level and Age [b]	QALYs Gained. *N*[c]	Incremental QALYs Gained, [d]	Cost (Thousands), *$*[c]	Incremental Cost (Thousands), *$*[d]	ICER (Thousands), *$/QALY* [e]	Optimal Screening Strateg[f]
Severe comorbidity						
76 y						
FIT	10.1	10.1	395	395	39	
FSIG	20.8	10.8	930	535	50	
COL	25.7	4.8	1329	399	83	X
77 y						
FIT	9.1	9.1	425	425	47	
FSIG	18.2	9.1	995	571	63	
COL	22.4	4.1	1400	404	99	X
78 y						
FIT	8.1	8.1	460	460	57	
FSIG	15.3	7.2	1071	611	85	X
COL	18.6	3.3	1483	412	125	
79 y						
FIT	7.4	7.4	493	493	67	X
FSIG	13.6	6.1	1134	640	105	
COL	16.4	2.8	1554	420	150	
80 y						
FIT	6.7	6.7	528	528	79	X
FSIG	11.6	4.8	1200	672	140	
COL	13.9	2.3	1625	424	184	

Abbreviations: FIT = fecal immunochemical test; FSIG = flexible sigmoidoscopy; COL = colonoscopy; LY = life-year; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life-year.

[a] QALYs gained and costs per 1000 persons, discounted at 3% per year. Results are also displayed in **Figure 10.2**. Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease and in case of a history of acute myocardial infarction; as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis, or AIDS; and as having no comorbid conditions if none of these conditions are present. In elderly persons without previous screening with no, moderate, and severe comorbid conditions, none of the screening strategies are cost-effective from age 87, 84, and 81 y onwards, respectively (**Figure 10.1**).

[b] All screening strategies consist of a 1-time screening examination followed by diagnostic and surveillance colonoscopies if indicated.

[c] Compared with no screening.

[d] Compared with the next less effective and efficient strategy, which is no screening for the first screening strategy mentioned at each age.

[e] Incremental cost per incremental QALY gained.

[f] The most effective, still cost-effective screening strategy based on a threshold for the willingness to pay per QALY gained of $100 000.

[g] In elderly persons without previous screening with no comorbid conditions aged 76 to 78 y, FIT screening is dominated by a combination of sigmoidoscopy screening and no screening (**Figure 10.2**).

Supplementary Table 10.6 Results of Sensitivity Analyses for CRC Screening Indicated in Elderly Persons Without Previous Screening

Analysis	Age Up to Which Screening Should Be Considered	Screening Strategy Indicated, by Age														
		76y	77y	78y	79y	80y	81y	82y	83y	84y	85y	86y	87y	88y	89y	90y
No comorbidity																
Base case	86y	COL	COL	COL	COL	COL	COL	COL	COL	FSIG	FIT	FIT				
Utility loss from COL, FSIG, and complication×0.5	86y	COL	COL	COL	COL	COL	COL	COL	COL	COL	SIG	FIT				
Utility loss from COL, FSIG, and complication×2	86y	COL	COL	COL	COL	COL	COL	FSIG	FIT	FIT	FIT	FIT				
Utility loss LYs with continuing care stage I and II CRC = 0.12 and 0.18, respectively	84y	COL	COL	COL	COL	COL	COL	COL	SIG	SIG						
Costs of COL, FSIG, and FIT×1.25	86y	COL	COL	COL	COL	COL	COL	COL	SIG	FIT	FIT	FIT				
Costs of COL, FSIG, and FIT×0.75	86y	COL	COL	COL	COL	COL	COL	COL	COL	COL	SIG	FIT				
Costs of CRC care×1.25	86y	COL	COL	COL	COL	COL	COL	COL	COL	SIG	FIT	FIT				
Costs of CRC care×0.75	87y	COL	COL	COL	COL	COL	COL	COL	COL	SIG	FIT	FIT	FIT			
Miss rates for COL and FSIG×2	86y	COL	COL	COL	COL	COL	COL	COL	COL	FIT	FIT	FIT				
No surveillance in adenoma patients	86y	COL	COL	COL	COL	COL	COL	COL	COL	COL	SIG	FIT				
CRC risk×1.25	86y	COL	COL	COL	COL	COL	COL	COL	COL	COL	COL	FIT				
CRC risk×0.75	86y	COL	COL	COL	COL	COL	SIG	SIG	SIG	FIT	FIT	FIT				
2 annual FITs as an additional screening strategy	86y	COL	COL	COL	COL	COL	COL	COL	COL	2 FITs	2 FITs	FIT				
Willingness to pay per QALY gained = $50 000	84y	COL	COL	COL	COL	COL	SIG	SIG	FIT	FIT						
Moderate comorbidity																
Base case	83	COL	COL	COL		SIG	FIT	FIT								
Utility loss from COL, FSIG, and complication×0.5	84	COL	COL	COL		COL	SIG	FIT	FIT							
Utility loss from COL, FSIG, and complication×2	83	COL	COL	COL		FIT	FIT	FIT								
Utility loss LYs with continuing care stage I and II CRC = 0.12 and 0.18, respectively	81	COL	COL	COL		SIG										
Costs of COL, FSIG, and FIT×1.25	83	COL	COL	COL		FIT	FIT	FIT								

Supplementary Table 10.6 Results of Sensitivity Analyses for CRC Screening Indicated in Elderly Persons Without Previous Screening (continued)

Analysis	Age Up to Which Screening Should Be Considered	Screening Strategy Indicated, by Age														
		76y	77y	78y	79y	80y	81y	82y	83y	84y	85y	86y	87y	88y	89y	90y
Costs of COL, FSIG, and FIT×0.75	84	COL	COL	COL		COL	SIG	FIT	FIT							
Costs of CRC care×1.25	83	COL	COL	COL		SIG	FIT	FIT								
Costs of CRC care×0.75	84	COL	COL	COL		SIG	FIT	FIT	FIT							
Miss rates for COL and FSIG×2	83	COL	COL	COL		SIG	FIT	FIT								
No surveillance in adenoma patients	84	COL	COL	COL		COL	COL	FIT	FIT							
CRC risk×1.25	83	COL	COL	COL		COL	COL	FIT								
CRC risk×0.75	83	COL	COL	SIG		FIT	FIT	FIT								
2 annual FITs as an additional screening strategy	83	COL	COL	COL		SIG	2 FITs	FIT								
Willingness to pay per QALY gained = $50 000	80	COL	COL	SIG												
Severe comorbidity																
Base case	80	COL	COL	SIG												
Utility loss from COL, FSIG, and complication×0.5	80	COL	COL	COL												
Utility loss from COL, FSIG, and complication×2	80	SIG	FIT	FIT												
Utility loss LYs with continuing care stage I and II CRC = 0.12 and 0.18, respectively	78	COL	SIG	SIG												
Costs of COL, FSIG, and FIT×1.25	80	SIG	SIG	FIT												
Costs of COL, FSIG, and FIT×0.75	80	COL	COL	COL												
Costs of CRC care×1.25	80	COL	COL	SIG												
Costs of CRC care×0.75	81	COL	COL	SIG		FIT										
Miss rates for COL and FSIG×2	80	COL	COL	FIT												
No surveillance in adenoma patients	81	COL	COL	COL		FIT										

Supplementary Table 10.6 Results of Sensitivity Analyses for CRC Screening Indicated in Elderly Persons Without Previous Screening (continued)

Analysis	Age Up to Which Screening Should Be Considered	Screening Strategy Indicated, by Age														
		76y	77y	78y	79y	80y	81y	82y	83y	84y	85y	86y	87y	88y	89y	90y
CRC risk×1.25	80	COL	COL	COL												
CRC risk×0.75	80	SIG	SIG	FIT												
2 annual FITs as an additional screening strategy	80	COL	COL	2 FITs												
Willingness to pay per QALY gained = $50 000	77	SIG	FIT													

Abbreviations: COL = 1-time colonoscopy screening; CRC = colorectal cancer; FIT = 1-time fecal immunochemical test screening; LY = life-year; QALY = quality-adjusted life-year; FSIG = 1-time sigmoidoscopy screening; 2 FITs = fecal immunochemical test screening during 2 consecutive years.

[a] Persons are classified as having moderate comorbid conditions if they have an ulcer, rheumatologic disease, peripheral vascular disease, diabetes, paralysis, or cerebrovascular disease, and in case of a history of acute myocardial infarction; as having severe comorbid conditions if they have chronic obstructive pulmonary disease, congestive heart failure, moderate or severe liver disease, chronic renal failure, dementia, cirrhosis and chronic hepatitis or AIDS; and as having no comorbid conditions if none of these conditions are present

Chapter 11

Cost-effectiveness of recommended surveillance for patients with colorectal adenomas

Article to be submitted:

RG Meester, I Lansdorp-Vogelaar, U Ladabaum

ABSTRACT

IMPORTANCE: A substantial amount of endoscopic resources are devoted to surveillance of patients with colorectal adenomas, while there is limited direct evidence to support current practice. The European Polyp Surveillance (EPoS) study is planned to evaluate currently recommended surveillance, but results are not expected for >10 years.

OBJECTIVE: To simulate lifetime colorectal cancer benefits and costs of currently recommended surveillance compared to less intensive surveillance or screening

DESIGN: Microsimulation screening analysis (MISCAN) model validated for surveillance evaluation

SETTING: Simulated U.S. population cohort

PARTICIPANTS: 50 year-old patients with adenomas detected through colonoscopy or fecal immunochemical test (FIT) screening

EXPOSURES: Patients with 1-2 small tubular adenomas (low-risk adenoma, LRA) were simulated to receive colonoscopy after 5 or 10 years; patients with 3-10 small tubular or ≥1 advanced adenomas (high-risk adenoma, HRA) received colonoscopy after 3 or 5 years. Reference outcomes were those of continued screening with colonoscopy or FIT.

MAIN OUTCOMES: Estimated lifetime colorectal cancer mortality; life-years and cost per life-year gained, discounted at 3% per year.

RESULTS: For adenoma patients identified through colonoscopy screening, continued screening was associated with an estimated lifetime colorectal cancer mortality risk of 18.8 per 1000 for patients with LRA, and 31.0 per 1000 for patients with HRA. Low-intensity colonoscopy surveillance was associated with decreased mortality risks of 15.7 and 22.8 per 1000, respectively (life-years gained 6 and 30 per 1000; cost/life-year US$ 4856 and 5752). High-intensity surveillance was associated with further decreased risks of 11.7 and 18.2 (life-years gained 26 and 58; incremental cost/life-year US$ 19,754 and 9125). In adenoma patients identified through FIT, baseline risks with continued screening were higher, however, surveillance outcomes were similar.

CONCLUSIONS: Microsimulation modeling suggests that currently recommended surveillance is highly effective and cost-effective in the long run compared to less intensive surveillance strategies. Evidence from the EPoS trial is needed to inform practice on more intermediate term effects, and for validation of long-term model projections.

INTRODUCTION

There is wide consensus across different societies regarding the merits of screening of average-risk patients aged 50-75 years.[114,115,246] Less clarity exists on the appropriate management of patients with removed adenomas. Although European and American guidelines largely agree on risk classifications of adenoma patients and recommended intervals for repeat examination,[323] there are surprisingly little cancer outcome data to support these recommendations. Long-term follow-up data from the National Polyp Study suggests that colonoscopy surveillance may reduce mortality by 50% compared to mortality in the general population.[193] However, concerns have been raised that intensive surveillance may no longer be cost-effective with improvements in the quality of colonoscopy.[141] Some cost-effectiveness models have suggested that currently recommended surveillance for patients with low-risk adenomas may overuse resources.[324,325] The European Polyp Surveillance (EPoS) trial is planned to study the question of appropriate surveillance of patients with resected adenomas, but results are not expected for more than a decade.[146]

To inform policy makers on the appropriate intensity of surveillance in patients with low-risk adenoma (LRA) and high-risk adenoma (HRA) surveillances, we used an established microsimulation model to estimate benefits and costs associated with currently recommended surveillance versus less intensive surveillance strategies.

METHODS

The MISCAN-Colon model (**Chapter 2**) was used to evaluate screening and surveillance intervention strategies in a virtual population similar to the United States in terms of life expectancy and colorectal cancer risk. We defined findings of 1-2 small (1-9mm) adenomas during screening as LRA, 3-10 small adenomas or ≥1 larger adenoma as HRA, and advanced adenomas as lesions >9mm in diameter, with >25% villous component, or with high-grade dysplasia. We did not model histological adenoma features, only size and multiplicity.

Effectiveness of screening and surveillance

The effects of screening and surveillance for colorectal cancer follow from the model's natural history assumptions and the test's assumed ability to detect and remove precursor adenomas and cancer (see Chapter 9 for test performance assumptions). Modeled effects of fecal immunochemical test (FIT) and colonoscopy screening have been shown to be consistent with observational data.[105,107] Overall estimated 20-year colorectal cancer mortality risk during surveillance is concordant with National Polyp

Study data (**Figure 2.6**).[326] For the present study, we also compared modeled rates of adenoma detection in surveillance to reported rates from the literature (**Supplementary Figure 11.1-4**, **Supplementary Table 11.1**). In addition to our default model for the United States, we evaluated three alternative model variants.

The default model was consistent with observed adenoma detection rates (ADRs) for patients with LRA and HRA,[54,327,328] and with advanced adenoma detection rates (A-ADRs) for patients with LRA,[54,55,142,327,328] but underestimated the A-ADR for HRA patients.[54,55,327-329] Alternative models with assumed higher adenoma miss rates or adenoma growth rates matched better with observed A-ADRs for HRA, but overestimated A-ADRs in LRA patients. Main study outcomes (described below) are reported for the default model, as well as two of the alternative models (as sensitivity analyses).

Risk of complications

Assumed age-specific complication risks for colonoscopy use were based on published literature.[271,272] Overall complication rates increased exponentially from 4.1 per 1000 for age 50 years to 146.1 per 1000 for age 100 years.

Screening and treatment cost assumptions

Screening, complication, and treatment costs were approximated using 2007 or older Medicare payment and co-payment rates (treatment cost data were from 1998-2003),[152,277,278] and included patient time valued at median US wage.[279] Costs were updated to 2015 using the general consumer price index.[279]

Outcomes

The main study outcomes were lifetime colorectal cancer mortality, life-years, and cost per life-year gained, according to baseline adenoma risk category and primary screening method. Life-years and costs were discounted at the conventional annual rate of 3%.[38] We defined costs per life-years of <US $50,000 as very cost-effective, and ratios exceeding US $100,000 as not cost-effective. Other outcomes reported include cancer incidence and the endoscopy resources used for surveillance and screening.

Analysis

We simulated 10 million men and women born Jan 1st 1965. Patients were screened at age 50 years with either colonoscopy or FIT as the primary test. For the main analysis, we included patients with adenomas detected in screening, which we classified as LRA or HRA according to above definitions. For each risk class, we evaluated a less and more intensive scenario for colonoscopy surveillance analogous to the strategies planned for evaluation in the EPoS study: LRA patients received

examination at 5 or 10 years; HRA patients were examined at 3 or 5 years. Successive surveillance intervals were determined based on the latest findings and – in case of a negative findings – findings during previous examinations (**Supplementary Table 11.1**). Surveillance was continued up to the first of several events: age 85 years, diagnosis of cancer, or death. Outcomes were compared across the two scenarios and against a scenario in which all patients returned to screening in 10 years after the index examination.

Sensitivity analysis

Several alternative scenarios to the base-case were evaluated for their influence on cost-effectiveness of surveillance: a stopping age for screening of 80 years instead of 75 years; indefinite surveillance up to age 100 years; surveillance of an older age-cohort of 70 year olds with and without prior screening; faster assumed adenoma growth (no changes in cancer risk); lower assumed colonoscopy quality (-10% sensitivity for all lesions); higher assumed colonoscopy quality (98% sensitivity for all lesions, and 98% exam completeness); 50% higher endoscopy cost; and, 50% lower treatment cost.

RESULTS

With colonoscopy screening offered to patients of age 50 years, 2.0 million patients (21.9%) had detected adenomas. Of these, 1.6 million patients were classified LRA (17.3%), and 0.4 million (4.6%) were HRA. In FIT screening, 453 thousand patients (4.8%) had a positive test result at age 50, 167 thousand patients (1.8%) had adenomas detected during follow-up colonoscopy, 93 thousand (1.0%) had LRA, and 74 thousand (0.8%) had HRA.

Cancer outcomes

In LRA patients identified through colonoscopy screening, continued screening was associated in the model with a lifetime colorectal cancer risk of 64.4 per 1000, and a lifetime cancer mortality risk of 18.8 per 1000 (**Table 11.1**, **Figure 11.1**, **Figure 11.2**). Low-intensity surveillance, decreased cancer incidence and mortality risks to 59.1 and 15.7 per 1000, respectively, a reduction of 8% and 17% compared to no surveillance. High-intensity surveillance further decreased the risks to 48.5 (-25%, compared to no surveillance) and 11.7 per 1000 (-38%).

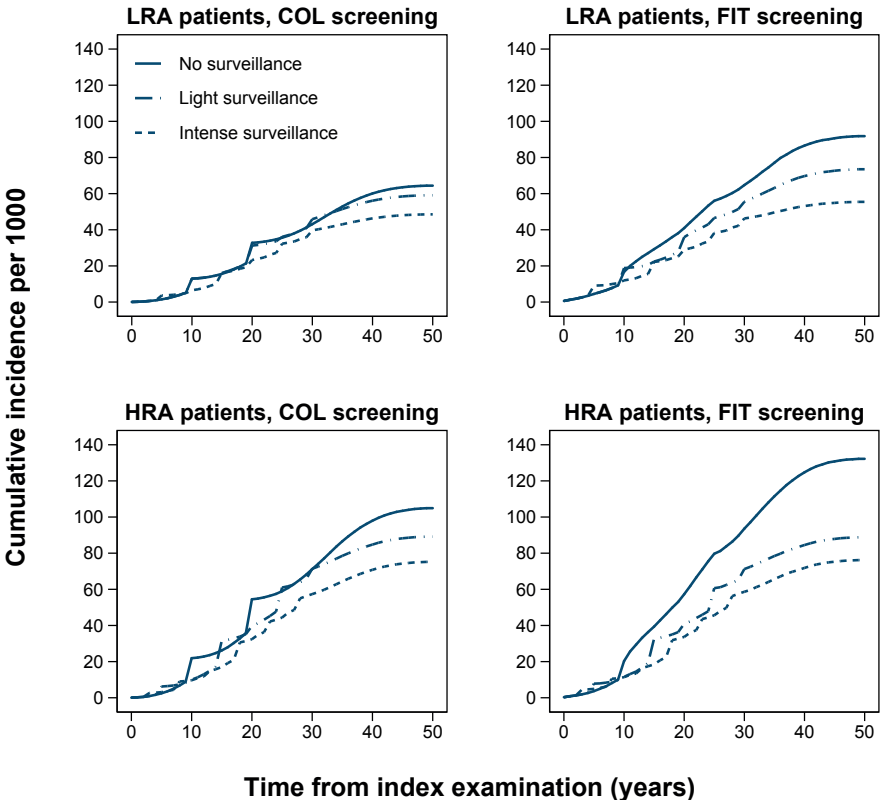


Figure 11.1a-d CRC incidence among adenoma patients in a colonoscopy (**left**) and FIT (**right**) screening setting with surveillance.

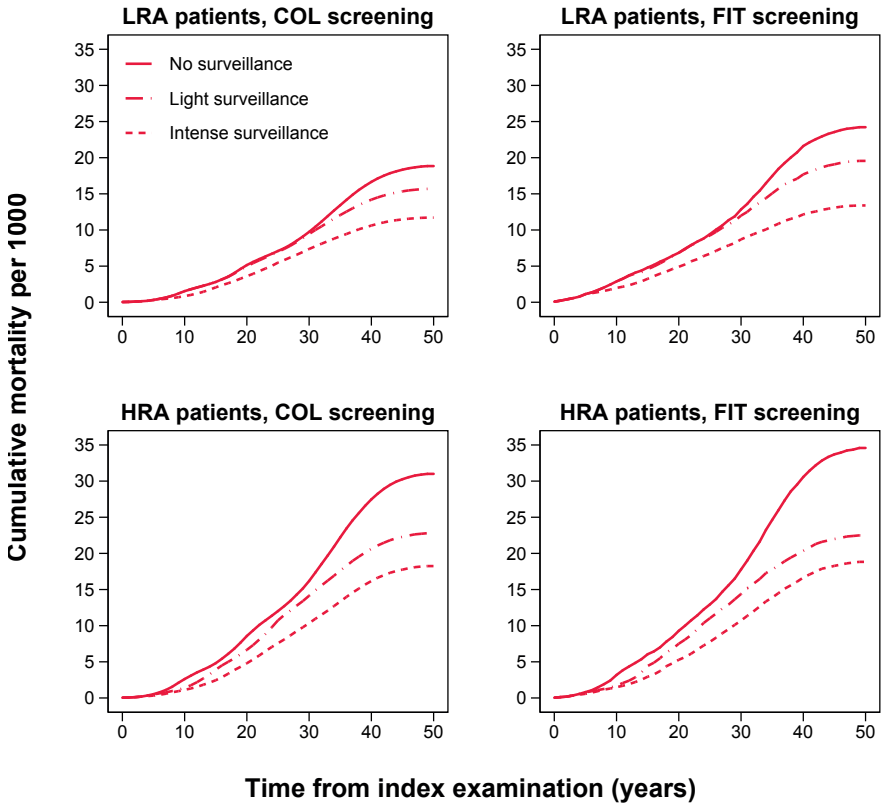


Figure 11.2a-d CRC mortality among adenoma patients in a colonoscopy (**left**) and FIT (**right**) screening setting with surveillance.

In LRA patients identified through FIT screening, lifetime cancer incidence and mortality risks with continued screening were higher than the estimated risks for colonoscopy screening, 91.8 and 24.2 per 1000, respectively (**Table 11.1**, **Figure 11.1**, **Figure 11.2**). Surveillance had stronger outcome effects, with similar estimated risks resulting as were found for colonoscopy screening. With low-intensity surveillance, incidence was 73.5 (-20%) and mortality was 19.6 per 1000 (-19%); with high-intensity surveillance the risks were 55.4 (-40%) and 13.4 (-45%), respectively.

HRA patients identified through colonoscopy screening had an absolute lifetime incidence risk of 104.9 per 1000 with no surveillance, and a mortality risk of 31.0 per 1000 (**Table 11.1**, **Figure 11.1**, **Figure 11.2**). Light-intensity surveillance reduced estimated risks to 89.2 (-15%) and 22.8 (-26%) per 1000, respectively, and high-intensity surveillance further reduced estimated risks to 75.2 (-28%) and 18.2 (-41%) per 1000.

Table 11.1 Simulated effectiveness and cost of surveillance with background colonoscopy and FIT screening.

		COL screening			FIT screening		
		Survellance intensity			Survellance intensity		
Risk group	Outcome	None	Low	High	None	Low	High
LRA	Incidence	64.4	59.1	48.5	91.8	73.5	55.4
	Stage I-II	42.3	40.8	35.4	63.9	50.3	40.3
	Stage III-IV	22.1	18.2	13.2	27.9	23.2	15.2
	Mortality	18.8	15.7	11.7	24.2	19.6	13.4
	LY gained	0	6	26	0	10	41
	Colonoscopies	2737	3014	4178	1550	2593	4209
	Cost (US$ mln)	5.0	5.1	5.4	5.3	5.6	6.0
	Treatment	2.7	2.6	2.1	4.0	3.3	2.6
	Endoscopy	2.4	2.5	3.3	1.4	2.3	3.4
	ICER	0	4856	19754	0	22469	15811
HRA	Incidence	104.9	89.2	75.2	132.2	88.8	76.2
	Stage I-II	68.8	63.7	54.9	92.8	62.4	54.4
	Stage III-IV	36.1	25.5	20.3	39.4	26.4	21.8
	Mortality	31.0	22.8	18.2	34.6	22.5	18.8
	LY gained	0	30	58	0	42	68
	Colonoscopies	2718	3710	4895	1601	3706	4881
	Cost (US$ mln)	6.8	6.9	7.2	7.0	7.1	7.4
	Treatment	4.4	3.8	3.2	5.6	3.9	3.3
	Endoscopy	2.4	3.1	4.0	1.4	3.1	4.1
	ICER	0	5752	9125	0	2891	10791

Abbreviations: LRA = low-risk adenoma; HRA = high-risk adenoma; LY = life-year; mln = million; ICER = incremental cost-effectiveness ratio.

In a FIT-based screening setting, HRA patients again had higher risks than comparable patients in a colonoscopy screening setting, and surveillance reduced risks to a more comparable level to risks with surveillance in colonoscopy-detected LRA patients. Estimated lifetime risk of colorectal cancer was 132.2 per 1000, and lifetime mortality risk was 34.6 per 1000 (**Table 11.1**, **Figure 11.1**, **Figure 11.2**). Low-intensity surveillance reduced incidence and mortality already by a third to 88.8 (-33%) and 22.5 per 1000 (-35%), respectively, and intensive surveillance reduced incidence to 76.2 per 1000 (-42%) and mortality by almost half to 18.8 per 1000 (-46%).

The risk reductions associated with surveillance translated to a gain of 6-10 life-years with low-intensity surveillance and 26-41 with high-intensity surveillance for 1000 patients with LRA. For patients with HRA, benefits were larger: 30-42 for low-intensity surveillance and 58-68 for high-intensity surveillance.

Resources

Colonoscopy screening was associated with a total number 2718-2737 colonoscopies per 1000 patients, compared to 1550-1601 with FIT-based screening (**Table 11.1**). Low-intensity surveillance increased use of colonoscopy resources to 2593-3014 for 1000 LRA patients and 3706-3710 per 1000 HRA patients. High-intensity surveillance increased colonoscopies to similar amounts for LRA (4178-4209) and HRA patients (4881-4895). Relative cost increases from no to intensive surveillance were smaller due to high averted cancer treatment costs (**Table 11.1**). For patients with LRA, total cost of screening and treatment varied from US$ 5.0-5.3 million with no surveillance to US$ 5.4-6.0 million with high-intensity surveillance (8-13% increase). For HRA patients, total costs varied even less, from 6.7-6.9 with no surveillance to 7.1-7.3 with high-intensity surveillance (6% increase).

Cost-effectiveness

Cost-effectiveness ratios were lower in general for HRA patients than LRA patients, and varied by primary screening method. For LRA patients, incremental cost-effectiveness ratios for low- and high-intensity surveillance were US$ 4856-22,469 and US$ 19,754-15,811, respectively (**Table 11.1**, **Figure 11.3**). For HRA patients, incremental cost-effectiveness ratios for low-intensity surveillance were US$ 2891-5752, and for high-intensity surveillance ratios varied US$ 9125-10,791.

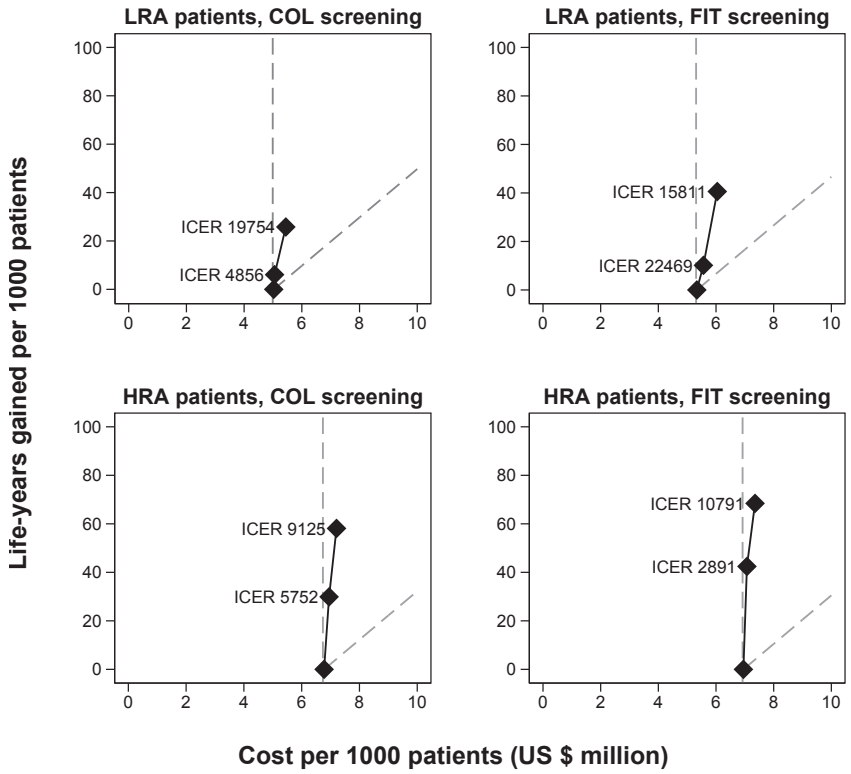


Figure 11.3a-d Benefits and costs of surveillance among adenoma patients in a colonoscopy (**left**) and FIT (**right**) screening setting. [a]

Abbreviations: LRA = low-risk adenoma; HRA = High-risk adenoma.

[a] The vertical dashed lines represent cost-saving thresholds for surveillance, and the diagonal dashed lines represent US $ 100,000 cost per life-year thresholds relative to screening. Screening (bottom), low-intensity surveillance (middle), and high-intensity surveillance (top) scenarios are represented by the black diamonds.

Sensitivity analysis

Low- and high-intensity surveillance remained cost-effective for scenarios evaluated in sensitivity analyses (**Table 11.2, Supplementary Table 11.3**). Although relative surveillance benefits decreased if screening was extended up to older age (1-33 life-years gained for LRA patients, and 17-56 for HRA patients), surveillance was initiated at an older age (14-38 for LRA, and 34-70 for HRA), large adenoma prevalence was increased without changes to overall cancer incidence (8-38 for LRA, and 21-51 for HRA), and assumed colonoscopy sensitivity was higher (5-30 for LRA, and 26-61 for HRA), estimated incremental cost-effectiveness ratios were nowhere higher than US$ 67,895.

Table 11.2 Sensitivity analysis of the benefits and cost of surveillance against colonoscopy screening.[a]

Risk group	analysis	Cost (US$ million)			Life-years gained			Cost/Life-year (US$)		
		Surveillance intensity								
		None	Low	High	None	Low	High	None	Low	High
LRA	Main	5.0	5.1	5.4	0	6	26	-	4856	19754
	End age screening (80y)	5.2	5.1	5.4	0	1	17	-	-10614	18295
	End age surveillance (100y)	5.0	5.1	5.6	0	6	27	-	10127	25117
	Older unscreened pts (70y)	2.6	3.0	3.1	0	14	20	-	23636	30799
	Older screened pts (70y)	3.1	2.8	2.8	0	31	38	-	-8297	-917
	Fast adenoma growth [b]	4.7	4.7	5.1	0	7	22	-	5030	28975
	Low quality (-10% sensitivity)	5.4	5.5	5.8	0	7	30	-	5650	15806
	High quality (98% sensitivity)	4.4	4.4	4.9	0	5	19	-	6181	33416
	High endoscopy cost (+50%)	6.2	6.3	7.1	0	6	26	-	16581	40967
	Lower treatment cost (-50%)	3.7	3.8	4.4	0	6	26	-	14153	31090
HRA	Main	6.8	6.9	7.2	0	30	58	-	5752	9125
	End age screening (80y)	6.8	6.9	7.2	0	17	45	-	6620	9125
	End age surveillance (100y)	6.8	7.1	7.4	0	32	62	-	10536	9310
	Older unscreened pts (70y)	3.9	4.1	4.3	0	34	49	-	4040	19542
	Older screened pts (70y)	4.7	3.3	3.7	0	48	59	-	-28960	38424
	Fast adenoma growth [b]	5.6	5.8	6.2	0	21	39	-	13660	21897
	Low quality (-10% sensitivity)	7.2	7.4	7.7	0	34	64	-	5379	7364
	High quality (98% sensitivity)	6.1	6.2	6.6	0	26	50	-	6087	14741
	High endoscopy cost (+50%)	8.0	8.5	9.2	0	30	58	-	17939	25526
	Lower treatment cost (-50%)	4.6	5.0	5.6	0	30	58	-	15063	20964

Abbreviations: LRA = low-risk adenoma; HRA = high-risk adenoma; pts = patients.

[a] Costs and life-years presented are per 1000 patients. Results for FIT screening are in **Supplementary Table 11.3**.

[b] We assumed a longer mean duration from adenoma onset to large adenoma size, but similar overall adenoma dwell time.

DISCUSSION

In regard of the lack of evidence for long-term effectiveness and cost-effectiveness of currently recommended surveillance for patients with colorectal adenomas, we validated an established microsimulation screening analysis model to assess benefits and costs of various colonoscopy surveillance strategies. The results of our study suggest that currently recommended surveillance may be very cost-effective in the long term compared to less intensive surveillance or continued screening. These findings underscore the appropriateness of current professional guidelines in the United States, which recommend examination of patients with low-risk adenomas with colonoscopy after 5 years and of patients with high-risk adenomas after 3 years.

In our study, patients with adenomas had high estimated risks of developing cancer compared to the general population in the United States: 6-9% for patients with LRA and 10-13% for patients with HRA without surveillance, versus 4.5% for an average individual.[330] Surveillance was estimated to decrease absolute risks by up to 3% compared to colonoscopy screening, and by up to 5.6% compared to FIT-based screening. Corresponding benefits in terms of life-years gained were maximum 68 per 1000 patients with discounting (162 without discounting), which is substantial in comparison with other health care interventions.[29] In contrast, the increase in the cost of care was only modest, with no more than 13% higher costs. Averted treatment costs weighed more heavily on total expenditure than the additional colonoscopies, which explains the relative low cost-effectiveness ratios of less than US$ 25,000 per life-year gained, well below the often applied threshold of US $50,000 or the GDP-linked WHO-CHOICE guideline (3xGDP).

We found that surveillance was associated with similar estimated risks for adenoma patients identified through FIT and colonoscopy screening. However, compared to FIT-based testing, surveillance benefits were relatively higher. This suggests that FIT screening is inadequate for patients with adenomas, and that there may be a more pressing need for surveillance in FIT-based screening settings than in colonoscopy screening settings. There were other factors which influenced cost-effectiveness of surveillance, including a higher screening cessation age of 80 years, higher colonoscopy quality, and older age. Surveillance remained cost-effective compared to screening (ratios <US$ 70,000) for all sensitivity analyses.

In the EPoS study design, it was assumed that less intensive surveillance is non-inferior to more intensive surveillance in terms of the incidence reduction, with 25% (LRA) and 35% (HRA) difference thresholds for inferiority.[146] We found that high-intensity surveillance was associated with substantial reductions in incidence and mortality compared to low-intensity, although differences generally remained below the above thresholds. The apparent conflict in results is due to the different study horizons: while we evaluated lifelong effectiveness and costs of surveillance, EPoS will look at effectiveness over a 10 year period. More frequent testing may initially increase diagnosis of early-stage cancer (**Figure 11.1**). In clinical studies such as EPoS, it is important to realize that benefits may accrue over longer periods of time, and that small effects in the short-term may not rule out more substantial long-term effects.

Our estimated relative mortality risks for LRA and HRA patients were consistent with data from the Norwegian cancer registry.[145] In a study by Løberg and colleagues (reflecting a situation without screening of average-risk patients), patients with detected LRA and HRA had standardized mortality ratios of 0.75 (95%CI, 0.63-0.88) and 1.16 (95% CI, 1.02-1.31), respectively, compared to the general population. Relative

to previous model estimates of the background colorectal cancer mortality risk in the United States (2.8%, without discounting), our base-case estimates of the relative risk for adenoma patients were 0.68 for LRA and 1.11 for HRA with colonoscopy every 10 years.[263] These findings suggests that surveillance normalizes cancer risk for adenoma patients.[107]

For this study, we also validated our model to published data on ADR during initial surveillance colonoscopy. Our model reproduced observed ADRs for both LRA and HRA patients (**Supplementary Figure 11.1**). The model underestimated A-ADRs for HRA patients compared to observed data. Some of the difference may be explained by the fact that we did not explicitly model villousness or severe dysplasia. It may also imply that the base-case model underestimated adenoma growth rates for patients with previous HRAs. We assumed that adenoma growth rates in a person are random for each lesion, which may not be valid for patients who had previous HRA. Alternative models with lower assumed colonoscopy quality and higher assumed adenoma growth rates matched better with observed data for HRA patients (**Supplementary Figure 11.2-3**). Despite non-trivial differences between the base-case and alternative models in terms of absolute surveillance benefits, surveillance was also effective and cost-effective for these alternative models (**Table 11.2**).

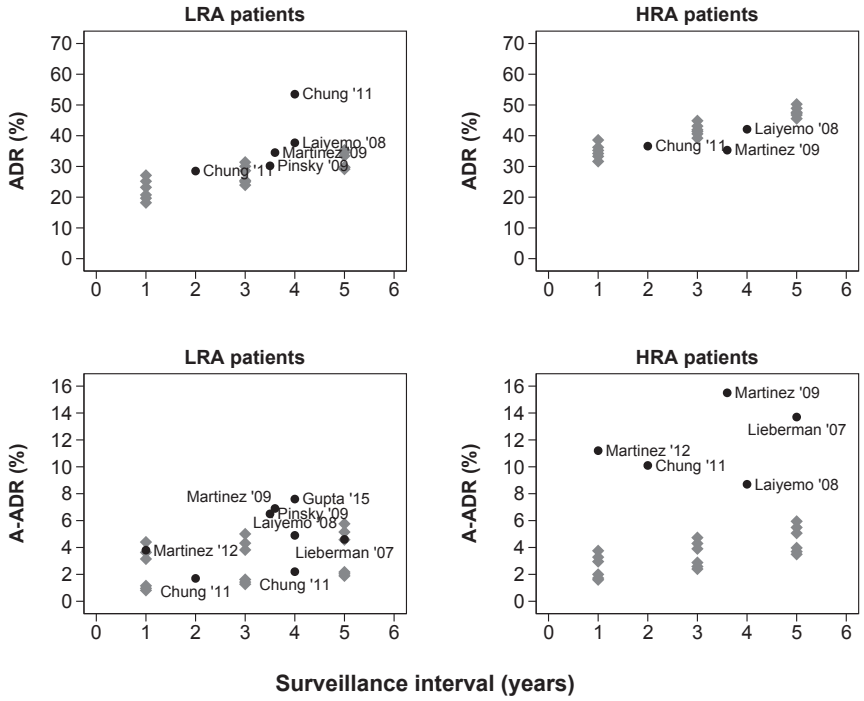
Our results partially contradict those from previous cost-effectiveness studies, which have suggested that surveillance after 5 years in patients with LRA may not be cost-effective.[324,325] The models used for those studies were relatively more simplistic, either not explicitly modeling the adenoma-carcinoma sequence or assuming fixed transtition rates from small to advanced adenomas and advanced adenomas to cancer, thereby not allowing for age dependency or person-specific risks of oncogenesis. Also none of the models were validated to data sources other than those used to inform the model. In our model, there is uncertainty regarding the adenoma dwelling time. Compared to other microsimulation models, the assumed adenoma dwell time is relatively short, which may have overstated surveillance benefits.[50]

There are other limitations for this study. First, our model does not explicitly describe adenoma histology, while surveillance guidelines also consider villousness, high-grade dysplasia, and serrated histology. However, observational do data suggest that multiplicity is the most important determinant of adenoma recurrence risk,[54] and that some of the other high-risk features may be strongly related to adenoma size.[53] Second, our analyses assumed 100% adherence with both screening and surveillance. In practice, screening adherence will be lower than 100%, and a surveillance indication may dramatically improve patient adherence relative to screening. Thus, in practice, surveillance benefits could be even higher. Finally, there is insufficient data to evaluate whether 1 year follow-up in patients with adenomas >20mm in di-

ameter or ≥5 adenomas at baseline is warranted.[59,60] In general, the option of further personalization in surveillance guidelines deserves more attention in future studies.

To conclude, we estimated that surveillance for adenoma patients as recommended by the United States Multi-Society Task Force is effective and cost-effective in the long term compared to less intensive surveillance or screening. Our findings support the use of colonoscopy surveillance for settings with sufficient colonoscopy capacity. Reductions in cancer-related mortality were also substantial for less intensive surveillance, which suggests that this could be considered as an alternative option for settings with a stronger aversion of or lower capacity for colonoscopy. Evidence from the EPoS study is needed to inform policymakers on the effects of surveillance in the intermediate-term, and to further refine the long-term model projections.

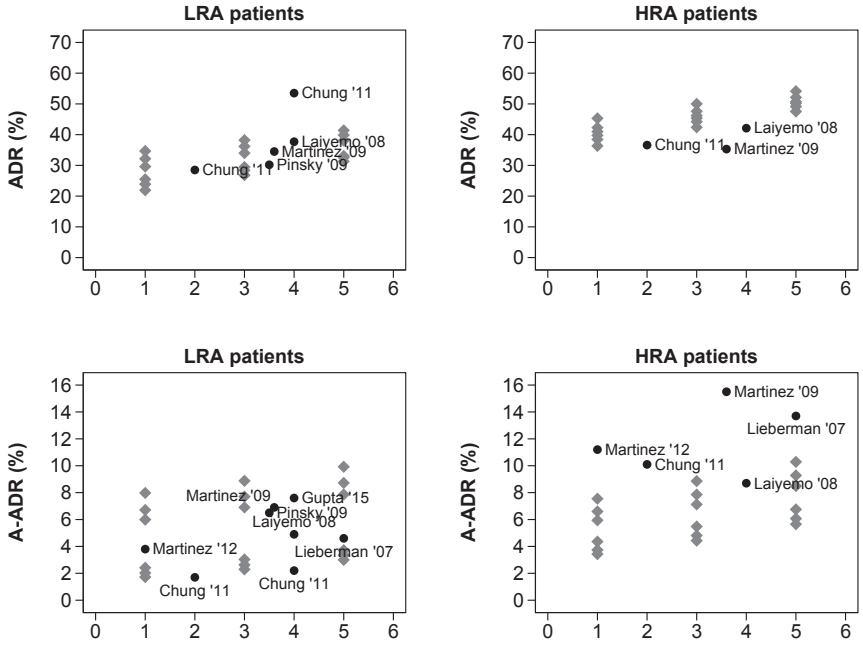
APPENDIX 11



Supplementary Figure 11.1a-d Simulated vs observed adenoma detection in patients with baseline LRA and HRA; base-case U.S. model. [a]

Abbreviations: ADR = Adenoma Detection Rate; A-ADR = Advanced Adenoma Detection Rate.

[a] Grey diamonds represent model results for 18 scenarios with varying age (mean: 55, 60, 65), screening test (FIT, colonoscopy), and surveillance interval (1y, 3y, 5y). Black dots represent observed data points from the literature. More details for these studies are provided in **Supplementary Table 11.1**.

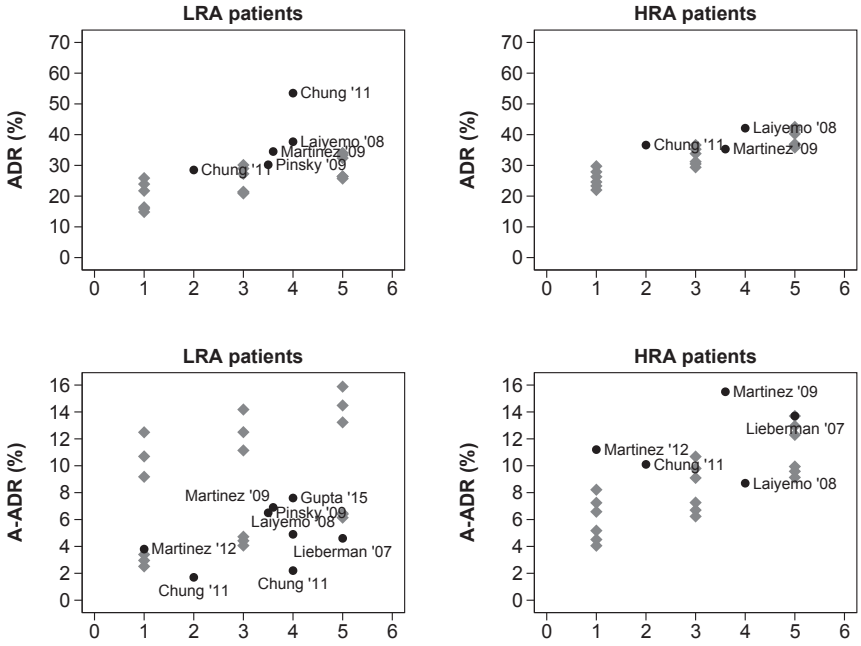


Supplementary Figure 11.2a-d Simulated versus observed adenoma detection in patients with baseline LRA and HRA; model with higher adenoma miss rates. [a, b]

Abbreviations: ADR = Adenoma Detection Rate; A-ADR = Advanced Adenoma Detection Rate.

[a] Grey diamonds represent model results for 18 scenarios with varying age (mean: 55, 60, 65), screening test (FIT, colonoscopy), and surveillance interval (1y, 3y, 5y). Black dots represent observed data points from the literature. More details for these studies are provided in **Supplementary Table 11.1**.

[b] In this model, we decreased assumed colonoscopy sensitivity by 5% for all lesions.

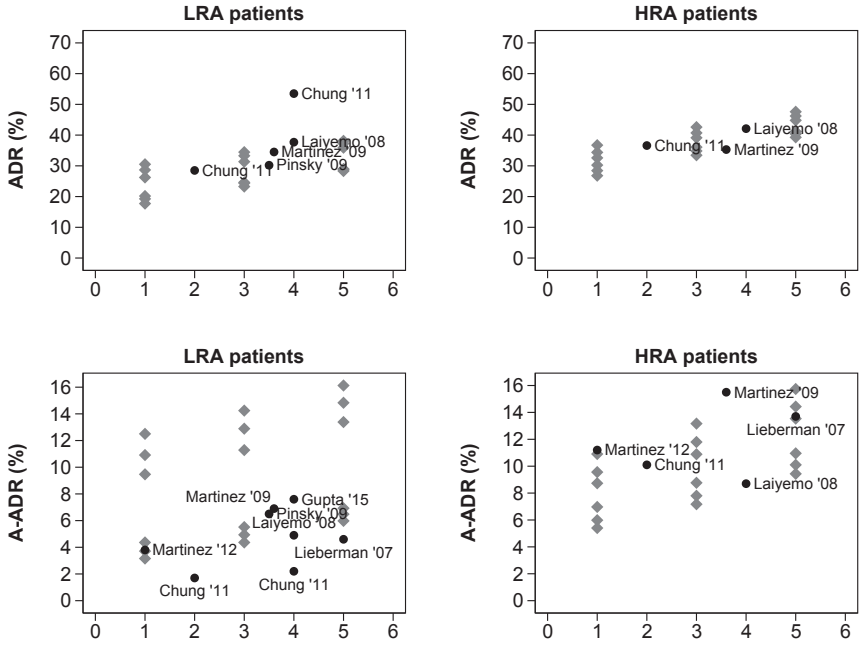


Supplementary Figure 11.3a-d Simulated versus observed adenoma detection in patients with baseline LRA and HRA; model with higher adenoma growth rates. [a, b]

Abbreviations: ADR = Adenoma Detection Rate; A-ADR = Advanced Adenoma Detection Rate.

[a] Grey diamonds represent model results for 18 scenarios with varying age (mean: 55, 60, 65), screening test (FIT, colonoscopy), and surveillance interval (1y, 3y, 5y). Black dots represent observed data points from the literature. More details for these studies are provided in **Supplementary Table 11.1**.

[b] In this model, we increased growth rates of adenomas >5mm in size.



Supplementary Figure 11.4a-d Simulated versus observed adenoma detection in patients with baseline LRA and HRA; model with higher adenoma miss and growth rates. [a, b]

Abbreviations: ADR = Adenoma Detection Rate; A-ADR = Advanced Adenoma Detection Rate.

[a] Grey diamonds represent model results for 18 scenarios with varying age (mean: 55, 60, 65), screening test (FIT, colonoscopy), and surveillance interval (1y, 3y, 5y). Black dots represent observed data points from the literature. Some more details for these studies are provided in **Supplementary Table 11.1**.

[b] In this model, we combined colonoscopy quality (model **Supplementary Figure 11.2**) with a higher adenoma growth rate for adenoma ≥5 mm.

Supplementary Table 11.1 Published adenoma detection rates during surveillance.

Index Finding	Reference	Patients, N	Age, mean	Interval, y	ADR, %	AADR, % [a]	Cancer %
LRA	Lieberman, 2007	473	63	2,3 or 5		4.6	
	Laiyemo, 2008	656	60.7	4	37.7	4.9	
	Pinsky, 2009	650	63	3.5	30.2	6.5	
	Martinez, 2009	4644	-	4	34.5	6.9	0.50
	Chung, 2011	671	57.8	2,3 or 5	45.8	2.4	
		355		2	28.5	1.7	
		316		3-5.5	53.5	2.2	
	Martinez, 2012	1194	-	1		3.8	
	Gupta, 2015	2477	62.3	3-5		7.6	
HRA	Lieberman, 2007	249	62-64	2,3 or 5		13.7	
	Laiyemo, 2008	855	61.5	4	42.1	8.7	
	Martinez, 2009	4523	-	4	35.3	15.5	0.80
	Chung, 2011	539	59.8	2,3	57.3	12.2	
		516		2	36.6	10.1	
	Martinez, 2012	2028	-	1		11.2	

Abbreviations: ADR = Adenoma Detection Rate; AADR = Advanced Adenoma Detection Rate.

Supplementary Table 11.2 Applied surveillance intervals. [a]

Index Finding	First Interval	First Finding	Second Interval	Second Finding	Third Finding
HRA	3 (5)	HRA	3 (5)	HRA	3 (5)
				LRA	5 (10)
				NA	5 (10)
		LRA	5 (10)	HRA	3 (5)
				LRA	5 (10)
				NA	10 (10-2S)
		NA	5 (10)	HRA	3 (5)
				LRA	5 (10)
				NA	10 (10-2S)
LRA	5 (10)	HRA	3 (5)	HRA	3 (5)
				LRA	5 (10)
				NA	5 (10)
		LRA	5 (10)	HRA	3 (5)
				LRA	5 (10)
				NA	10 (10-2S)
		NA	10 (10-2S)	HRA	3 (5)
				LRA	5 (10)
				NA	10 (10-2S)

Abbreviations: HRA = High-risk adenoma; LRA = low-risk adenoma; 10-2S= return to screening after 10 years.

[a] Intervals showed first are for the intense surveillance regimen; intervals for the light regimen are shown within parentheses

Supplementary Table 11.3 Sensitivity analysis of the benefits and cost of surveillance in a FIT screening setting.[a]

Risk group	Analysis	Cost (US $ million)			Life-years gained			Cost/Life-year (US $)		
		Surveillance Intensity			Surveillance Intensity			Surveillance Intensity		
		None	Low	High	None	Low	High	None	Low	High
LRA	Main	5.3	5.6	6.0	0	10	41	-	22469	15811
	End age screening (80y)	5.3	5.6	6.0	0	6	33	-	41850	18077
	End age surveillance (100y)	5.3	5.6	6.2	0	11	42	-	24653	19695
	Older unscreened pts (70y)	3.9	4.3	4.5	0	19	30	-	23248	19277
	Older screened pts (70y)	3.3	3.3	3.4	0	23	32	-	2204	16267
	Fast adenoma growth [b]	4.6	5.1	5.7	0	8	33	-	67895	22967
	Low quality (-10% sensitivity)	5.9	6.2	6.6	0	13	48	-	5379	7364
	High quality (98% sensitivity)	4.4	4.5	5.1	0	9	30	-	15997	27199
	High endoscopy cost (+50%)	6.0	6.7	7.7	0	10	41	-	67363	34277
	Lower treatment cost (-50%)	3.3	3.9	4.7	0	10	41	-	56128	26372
HRA	Main	7.0	7.1	7.4	0	42	68	-	2891	10791
	End age screening (80y)	6.9	7.1	7.4	0	30	56	-	6282	10791
	End age surveillance (100y)	7.0	7.2	7.6	0	45	73	-	6504	11254
	Older unscreened pts (70y)	4.7	4.9	5.1	0	44	60	-	3239	17551
	Older screened pts (70y)	4.4	4.0	4.2	0	55	70	-	-8078	15154
	Fast adenoma growth [b]	5.4	6.2	6.6	0	29	51	-	25888	18687
	Low quality (-10% sensitivity)	7.4	7.6	7.8	0	44	74	-	5262	6309
	High quality (98% sensitivity)	6.3	6.3	6.6	0	41	61	-	-1452	17516
	High endoscopy cost (+50%)	7.6	8.6	9.4	0	42	68	-	23862	28446
	Lower treatment cost (-50%)	4.2	5.1	5.7	0	42	68	-	22417	23050

Abbreviations: LRA = low-risk adenoma; HRA = high-risk adenoma; pts = patients.

[a] Costs and life-years presented are per 1000 patients.

[b] We assumed a shorter mean duration from adenoma onset to large adenoma size, but similar overall adenoma dwell time

PART V

General discussion

Colorectal cancer screening is a rapidly developing and expanding discipline in preventive medicine. In this thesis, we attempted to answer some of the current questions regarding the public health impact of screening (Part II, Chapter 3-5), the importance of various modifiable factors related to the quality of screening, including the test used for screening (Part III, Chapter 6-9), and the potential for more personalized screening or surveillance (Part IV, Chapter 10-12). For most of the questions addressed in this thesis no direct answers were available from empirical data at the time of study conduct. We therefore used microsimulation modeling to provide an initial hint of an answer in some cases, and in other cases, more definite answers. All studies and study findings are briefly summarized below. Subsequently, strengths, limitations, and implications for future practice and research will be addressed.

STUDY SUMMARIES

Potential public health impact of screening

In Part II, we estimated the public health impact of and capacity for potential higher uptake of screening in the United States. First, in Chapter 3, we estimated for the American Cancer Society what could be the impact of increased colorectal cancer screening on future cancer incidence and mortality. Their request for this analysis derived from the 2014 National Colorectal Cancer Roundtable initiative to increase colorectal cancer screening rates in the United States from the 2013 level of approximately 60% to 80% by 2018. To estimate the impact we simulated colorectal cancer test use as observed in the past from National Health Interview Survey data, and compared a scenario of stable future screening uptake to a scenario in which screening uptake would increase to 80% by 2018. Our results suggested that, if the United States succeeds to screen an additional 20% of the screening-eligible population, this could reduce disease incidence by 22% and mortality by 33%. The latter converted to a total of 280,000 avertible colorectal cancer cases and 200,000 avertable colorectal cancer deaths through 2030.

In Chapter 4, we assessed whether there actually would be sufficient colonoscopy capacity to screen 80% of the United States population. Microsimulation was used to estimate colonoscopy demand within a national screening program with FIT or colonoscopy as the primary screening method. A national survey (Survey of Endoscopic Capacity [SECAP]) was conducted to estimate colonoscopy volume in 2012, as well as the additional available capacity. The results suggested that, currently, there is sufficient national capacity to screen 80% of the population using FIT, colonoscopy, or a mix of the two tests. Screening was estimated to require

5-13 million colonoscopies depending on the primary screening test, while current estimated colonoscopy volume is 15 million. Sensitivity analyses pointed out that actual observed patterns of testing from National Health Interview Survey data are more intensive than recommended screening, with over 20 million estimated exams required with 80% screening uptake. However, the total estimated colonoscopy capacity of approximately 25 million would be sufficient to meet this higher need.

In Chapter 5, we looked beyond the potential impact of achieving 80% colorectal cancer screening by 2018, at the total potential impact of screening with full screening uptake. We expressed the maximum potential impact of screening on colorectal cancer mortality in terms of the population attributable fraction (PAF) of risk for nonuse of screening. PAF was derived in two ways: first, using microsimulation modeling, comparing disease-related mortality in a scenario with 100% hypothetical screening uptake to that recently observed in Surveillance Epidemiology and End Results data; second, we followed the traditional approach to estimating PAFs using Levin's formula, which was informed by empirical data on the relative risk reduction associated with screening use and the prevalence of screening. Our modeling suggested that full uptake of screening in the United States population could have reduced colorectal cancer mortality by 63% compared to the 2010 level. Our results also suggest that Levin's formula may underestimate the PAF compared to modeling, with a considerably more conservative estimate of 46% resulting for the PAF. The difference was mainly due to incorporation by the model of lagged effects of past increases in screening and disparities in effects by age.

Determinants of screening effectiveness

In Part III, we estimated the potential influence of several effect determinants on outcomes of colorectal cancer screening programs. First, in Chapter 6, we looked at the effect of adherence on the comparative benefits of colonoscopy versus a program of annually repeated stool-based testing for colorectal cancer. For this study, we used observed test adherence and outcome data from the National Colonoscopy Study, a multi-center randomized clinical trial of colonoscopy versus ≥4 rounds of annual high-sensitivity FOBT (Hemoccult Sensa, Beckman Coulter Inc.) in approximately 3500 average-risk persons. Modeling was used to project out the long-term risks of cancer incidence and mortality associated with observed adherence, and to contrast these with hypothetical scenarios of no screening and 100% screening adherence. In the National Colonoscopy Study, adherence was higher for colonoscopy screening (86%) than for sFOBT screening (50% completing all sFOBTs after 4 years, and 80% completing ≥1), With observed study participant adherence, colonoscopy was estimated to be significantly more effective than sFOBT, with relative effects on cancer incidence of -32% versus -1% and on cancer-related mortality of -59% versus

-43%. Compared to screening with 100% adherence, colonoscopy with observed adherence resulted in a moderate loss of 14% of the total potential mortality reduction, compared to the 33% for sFOBT.

In Chapter 7, we evaluated the importance and effect of variation in colonoscopy quality, as measured by a physician's adenoma detection rate (ADR). This work succeeded previous work by our research consortium which had already found a strong association between ADR and fatal interval cancer risk.[124] Using a subset of these data consisting of only screening colonoscopies (nearly 60,000), we informed our microsimulation model for estimating the long-term impact of variable ADRs on cancer outcomes, colonoscopy volume, risk, and overall treatment and screening costs in a colonoscopy screening setting. We found estimated differences of 50-60% in lifetime incidence and mortality between the lower and upper ADR quintiles. Unlike suggestions from literature, the estimated colonoscopy burden for surveillance of adenoma patients did not rise to unacceptable levels with higher ADRs (+4.6% per 5% point ADR increase), and the increase in risk of complications was modest compared to the estimated decrease in cancer risk (max +0.6/1000 complications vs -3.0/1000 cancer cases per 5% point ADR increase). Perhaps most importantly, our work suggested that higher ADRs may result in lower net costs of screening and treatment than lower ADRs due to averted treatments (-3.2% per 5% point ADR increase).

In Chapter 8, we expanded our work on the impact of varying levels of ADR to settings of fecal-based testing, where colonoscopy is used for follow-up of positive results. For various ADR quintile, we contrasted FIT screening outcomes with previously simulated colonoscopy screening outcomes. Our results suggested that both colonoscopy and FIT are sensitive to variable ADR. However, FIT screening may be affected less by suboptimal levels of ADR: in the lowest ADR quintile, the estimated relative risk of colorectal cancer mortality with FIT vs colonoscopy was 1.22 (95%CI, 0.92-1.75), while in the lowest ADR quintile the relative risk was significantly less than one 0.85 (95%CI, 0.82-0.93). The relative stability of FIT outcomes was due to our assumptions that FIT primarily detects advanced adenomas,[161] and that colonoscopy providers primarily miss smaller adenomas.[163]

In Chapter 9, we looked at the understudied question to the effect of longer time to diagnostic colonoscopy in persons with a positive colorectal cancer screening result. Since no appropriate data were available to inform our model on the effect of time to diagnostic follow-up, we used modeling alone to estimate the risk associated with longer time to follow-up. The effect of delayed diagnostic testing applies only to screening tests other than primary colonoscopy, and we focused specifically on the increasingly popular FIT (gFOBT and sDNA evaluated as sensitivity analyses). In our analysis, we compared health outcomes and cost for a program of annual FIT

between ages 50 and 75 years with follow-up of positive results after various time intervals of up to 12 months. Our results suggested that increasing time to diagnostic testing may result in proportional increases in cancer incidence and, particularly, mortality, and may decrease the overall benefits of screening in life-years gained by almost 10%. Screening cost was affected much less by time to follow-up.

Personalizing screening and surveillance

In Part IV, we evaluated the cost-effectiveness of taking into account more than just age for recommendations of screening. First, in Chapter 10, the central study question was whether and up to what age to screen previously unscreened elderly people over age 75 years. For this study in specific, we used microsimulation modeling with life-tables stratified by people's comorbidity status. Persons were classified as either having no comorbidities, moderate comorbidities, or severe comorbidities. We evaluated the merits of several tests, including colonoscopy, flexible sigmoidoscopy, and FIT up to age 90 years. The results of our study suggested that screening is often warranted well beyond age 75 years. With some variation in cost-effectiveness for the different types of evaluated screening tests (less invasive testing generally being cost-effective up to older ages than more invasive testing), screening was indicated up to age 83-86 years in persons without comorbidities, up to age 80-83 years in persons with moderate comorbidities, and up to age 77-80 in persons with severe comorbidities with a cost-effectiveness threshold of US$ 100,000.

Finally, in Chapter 11, we evaluated the appropriate intensity of surveillance in patients with removed colorectal adenomas using modeling. We distinguished patients with low-risk (1-2 small tubular) and high-risk (3-10 small tubular, 1 or more larger) adenomas. Evaluated surveillance strategies were based on the European Polyp Surveillance (EPoS) study design:[146] patients with low-risk adenomas received colonoscopy at 5 or 10 years, and patients with high-risk adenoma at 3 or 5 years. Outcomes were compared to those of screening as recommended for average-risk patients with colonoscopy every 10 years or annual FIT. For this study specifically, we validated our model to published data on adenoma detection rates (any and advanced only) during surveillance examinations. The results of our study suggested that currently recommended, intensive surveillance is effective and cost-effective compared to less intensive surveillance or regular screening. In the base-case analysis, we found stable differences in mortality compared to no surveillance of relatively 38-46%, and incremental costs per life-year gained were generally below US $25,000.

METHODOLOGICAL CONSIDERATIONS

The practical implications of this thesis are as multi-faceted as its scope. As for any research project, the implications of our work are defined to large extent by its strengths and limitations. In the following paragraphs, we highlight some of the main strengths and limitations.[331]

Uncertainty of the estimated health impact of colorectal cancer screening

There were several factors which contributed to quality and reliability of our work for the American Cancer Society in Chapter 3, and the related study in Chapter 5. The Microsimulation Screening Analysis model is a well-established tool to assess the benefits of screening. Its predictions for the public health impact of colorectal cancer screening in the United States were derived from and underpinned by robust evidence from randomized controlled screening trials. It has been validated to multiple different studies,[190,192,263] and used for two consecutive decision analyses for the United States Preventive Services Task Force.[204,259] Our current analysis closely replicated the age and sex composition of the United States population and colorectal cancer test utilization. The analysis factored in demographic trends. Patterns of test utilization were based on observed data from National Health Interview Surveys, which is considered the principal source to assess the use of preventive health care services in the United States.[331]

Besides these strengths, there are also some notable limitations and uncertainties. The National Health Interview Survey data did not provide great detail about the different types of stool-based tests used for screening, while data suggest that performance varies by brand.[262,311] We assumed that people mainly used the common gFOBT (Hemoccult II, Beckman Coulter Inc.), and the increasingly popular FIT (OC Sensor, Eiken Chemical Inc.) after 2000 (only Chapter 3), which is a simplification of more complex reality. The data also did not allow us to tease out exactly which proportion of colorectal exams was performed for screening purposes, diagnostic purposes, and surveillance. A German study has suggested that health effects for diagnostic exams may be smaller than those for screening.[167] However, the size of the difference (91% vs 72-85% reduction) suggests that misrepresentation of the use of colonoscopy for symptoms may only have moderate impact on estimated test benefits. An underestimation of the use of exams for symptoms could have overstated effects of testing.

No trials have assessed the effect of colonoscopy on proximal disease. Although many observational studies have reported substantial effects,[165-167,332,333] some studies have reported no effect of colonoscopy for the proximal end of the colon.[244,334,335] We assumed no difference in the effect of colonoscopy proximal and distal to the

splenic flexure, apart from some effect of potential incompleteness of exams (i.e. not reaching the end of the colon in some persons). Analogous to colonoscopy, the current generation of stool-based tests (sFOBT, FIT, and multi-target stool DNA tests) have not been evaluated in any randomized controlled trials to this date. Cancer outcomes data are scant, generally, for stool-based tests, although some initial observational data have been published recently.[105,106] Our model predictions for colonoscopy and FIT efficacy are in line with currently available observational data for the effects of screening.[105,107] Although this suggest that the results of our model may be realistic, the evidence is relatively weak compared to evidence from randomized clinical trials.

We did not assess in Chapter 3 and 5 the population-level effect of screening on life-years gained (or QALY), which is more uncertain than the effect of screening on cancer-related mortality. Except for the UK Flexible Sigmoidoscopy Study Randomized controlled trials,[117] no trials have found significant reductions in all-cause mortality. Several studies reported even higher mortality in the screening group than the intervention group (not statistically significant).[120,162,188,189,265] The difficulty of observing all-cause mortality effects stems from the small proportion of total deaths attributable to colorectal cancer, which we mentioned to be only 1.3% in 2012 in our introductory chapter. The risk of colorectal adenomas and cancer may be associated with higher risk of other conditions, for example cardiovascular diseases,[336] which could mean that the effect on life-years may be smaller than we estimated for incidence and mortality. However, the fact that there are few very strong risk factors for colorectal cancer, suggests that this relation may be relatively weak.

Finally, we could not assess whether currently unscreened people have average-risk for colorectal cancer or competing causes of death. While higher risk for cancer would suggest that we have underestimated the potential public health impact of increased screening in the United States, higher competing risks or poorer overall health status would imply a lower benefit in terms of life-years. Since currently unscreened people are relatively less educated, insured, and are more often from ethnic minorities,[82,222] it is possible that our assumption has somewhat underestimated colorectal cancer benefits. Recent observational data suggest that the inverse association between screening uptake and other cause mortality may be relatively weak.[337]

Scratching the surface of the need versus capacity question

In contrast to the studies presented in Chapter 3 and 5, Chapter 4 on screening capacity in the United States was less prone to most of the limitations discussed above. Health outcomes were not considered. Correct classification of test indication was less important for our study since we evaluated hypothetical scenarios for future

screening, and as a sensitivity analysis, extrapolated test use patterns as observed in National Health Interview Survey data. As an indicator of the study's face-validity, both the survey and model estimates for colonoscopy volume in 2012 were similar to indirect estimates derived by combining estimates of population screening rates and population size[338]

A limitation of this work was that the Survey for Endoscopic Capacity in the United States did not allow for a very detailed assessment of colonoscopy volume and capacity by geographical sub-regions (beyond North, South, East, West). Capacity constraints are effective at the regional or even local level, given there is limited possibility for traveling to receive preventive services. Even if survey data were sufficiently detailed to look at regional or even local availability of services, it would have been unwieldy for the model to capture this level of detail. Local health systems should ultimately assess whether capacity is sufficient , and take potential measures for expansion where needed.

Other limitations of the study include a suboptimal response rate of approximately two third, and uncertainty regarding the physicians' estimates of potential available colonoscopy capacity on top of current estimated volume. Survey respondents were asked to estimate the number of additional colonoscopies that they could perform without additional resources. It is unknown whether these estimates reflect what could be done without changes to current practice or reflect a shifting of resources away from other procedures. Survey questions indicated that the main limiting factors to increasing capacity are resources (physicians, nurses, equipment).

Effectiveness indicators or modifiers

In the third section of this thesis, we studied more select population subgroups for the effectiveness of fecal versus endoscopic screening methods under various specific modifiable conditions. The same limitations as we discussed above for model predictions of population benefits of screening in general, apply to this section.

A key distinctive factor and strength for at least two of these studies, however, was the fact that we could use high-quality data to partially inform the studies and effect estimates. In stool-based testing persistent adherence is critical for favorable long-term benefits.[83,84,86,87,339] As discussed before in Chapter 1, long-term adherence and associated effects are unknown for sFOBT as well as FIT. The data we used from the National Colonoscopy Study (NCS) comprised adherence and outcome data for sFOBT across up to seven rounds of testing. This is the longest follow-up recorded in a trial for any of the current stool-based tests, after major trials from the nineties have studied less effective gFOBT tests. NCS was a multi-center randomized clinical trial comparing adherence and performance of colonoscopy versus sFOBT, safeguarding comparability of participants and services across both study arms. In our model,

NCS adherence and positivity data were closely replicated to assess corresponding long-term effects, and short-term predictions for advanced adenoma detection rates and cancer incidence could be validated against observed data. Given the similarity of our model to the model used to inform recent United States Preventive Services Task Force recommendations,[259] another strength of this study was we that we could immediately compare our predictions to those for the task force, to quantify the loss in colorectal cancer mortality benefits attributable to suboptimal adherence.

A limitation of this study was that the NCS was underpowered to statistically compare primary outcomes such as incidence and mortality. There were relatively few observations for study rounds 6-7, which mainly reflect data from only one study center. Finally, NCS generalizability may be limited to the extent that test other stool-test brands perform differently, participant navigation is not available in practice in some settings, and in the sense that persons included in the study gave consent a priori to be randomized to colonoscopy versus sFOBT. The latter means that the study cohort consists of persons who are in principal willing to undergo screening, which may be a very select group in some settings. The scope of the modeling analyses was limited to benefits, instead of also including harms and cost-effectiveness.

Similar to the adherence study of Chapter 6, the two studies on the impact of variation in adenoma detection of Chapter 7-8 also used high-quality data to populate the model. The dataset that was utilized is still the largest published to this date on the relationship between this principal quality indicator for colonoscopy and cancer outcomes. The complete dataset has been used to justify updating clinical quality guidelines in the United States (Kaiser Permanente Northern and Southern California).[124,131] We used a subset of nearly 60,000 screening-only exams from the total database, to look specifically at the relation between ADRs and interval cancer risk after screening colonoscopy. This allowed for an immediate comparison between the data and our model, which is particularly apt for simulating average-risk screening populations. The model closely matched data for four out of five ADR categories, which increased the likelihood that our assumptions represent actual practice variation.

Despite the relatively large dataset and number of cancers, there remains some uncertainty regarding the validity of ADR as a quality indicator. As the proportion of a physician's colonoscopies detecting one or more adenomas, ADR is an imperfect proxy for adequate management of all clinically relevant lesions. Although it was inversely related with interval cancer risk, confidence intervals for cancer incidence were still relatively wide, and no individual inter-quintile differences were statistically significant at a 95% confidence level. However, the broader empirical study by Corley and colleagues reported similar inverse relationships between ADR and outcomes that were statistically significant.

The ADR-outcome relationship was also studied only for pools of physicians according to ADR category. The precision of the quality indicator at the individual level could not be assessed, given cancer incidence is too low (~50 per 100,000 years). Being a product of both adenoma prevalence and the physician's ability to detect adenomas, it is unclear from what number of exams ADR differences truly reflect physician rather than case-mix differences: with a low number of 5-10 exams to measure ADR, differences will surely reflect mostly differences between examined individuals; even with 100 exams there is likely still substantial adenoma prevalence variation from the one to the next sample of persons (standard error ~5% point).[325] In our study, the number of included colonoscopy exams per physician varied substantially (median 375, range 77-1262), but it was not powered to assess the ADR-outcome relationship for physicians with a small included exam base.

Another general limitation of studies on ADRs and cancer outcomes is that the effect of improving ADRs has not been established. ADR was intended as a quality indicator for colonoscopy.[132] In some settings, however, it is treated more like a sufficient quality measure, being linked directly to physician payment modifiers to stimulate higher ADRs and better quality of care.[340] It is unclear whether such policy may induce gaming. Physicians may no longer adequately examine patients once the first polyp is removed, or may not adequately clear polyp margins, if ADR is the only intra-procedure quality metric used to assess colonoscopist performance. In this case, the estimated causal association between ADR and outcomes in our study would break down. Thus, an exclusive focus of managers and health care systems on ADRs for colonoscopy quality assessment may lead to a deterioration of colonoscopy quality and health outcomes rather than improvement.

Finally, we estimated costs of screening and treatment for the ADR study using Medicare payment and copayment rates from 2007 or earlier. These approximations may not reflect true costs, with payment rates from private insurers for screening tests known to be higher than those for Medicare enrollees.[341] Rates may also be outdated. Cancer treatment costs may have risen rapidly with the use of expansive new chemotherapeutic agents. We updated costs using the general consumer price index which may not reflect these increases in treatment costs. Our sensitivity analyses did suggest that colonoscopy cost assumptions were not critical for our overall conclusions.

The final chapter of Part III, Chapter 9, was different from the preceding chapters in the sense that no data were available to directly inform the model on the relationship between the outcome determinant and outcomes itself. The effect of time to diagnostic colonoscopy after positive fecal colorectal cancer tests is an understudied subject. It is known already for years that the interval to colonoscopy may vary substantially across screening participants.[342,343] Delays in the Veterans Affairs health care

system have aroused wide media attention in recent years.[344] Remarkably, apart from two small and underpowered empirical studies, no previous studies investigated the effect of delayed follow-up after positive stool-test results. Despite uncertainty on how sojourn time is distributed across cancer patients, our model showed that even with extreme assumptions (+50% average duration compared to base) there may be a clinically relevant effect of time to colonoscopy. By model parameter variation, we were further able to reveal that the positive predictive value of stool-based tests may be a crucial effect modifier. High-specificity tests such as low-sensitivity gFOBT, or less frequently repeated testing may greatly increase the risk of longer time to follow-up. Without modeling, it would have been difficult to establish these relationships.

A limitation of the study on time to diagnostic follow-up after positive fecal colorectal cancer test results is that the model assumed very regular (exponential) patterns for potential occurrence of cancer between a positive stool test result and follow-up. The exponential model for cancer progression basically assumes that cancer incidence and progression rates are constant over time. This may not represent real life. Although we varied the estimated mean duration in each cancer stage (sojourn times), we did not assess whether an altogether different time distribution would lead to different results. Empirical study is therefore needed to complement our work.

Bounded personalization

In Part IV of this thesis we studied two variations on the subject of personalized screening. The potential value of personalized medicine in general, and screening in particular, is substantial in theory. In practice, it is limited by availability of data to assess cancer and other-cause risk differences for a large amount of potential risk factors. We focused in this section on two questions, namely: under what comorbid conditions and up to what age elderly people should receive screening, if they did not receive any previous screening; and, what intensity of surveillance is appropriate for patients with classified low- or high-risk adenoma. Comorbidity classes and adenoma classes were determined previously by other studies. Available data were confined to those classes, which constrained the scope of these studies.

A key strength of the study on the potential benefits of screening in elderly unscreened persons is that it addressed a very practical and timely question for physicians in the United States. As we mentioned in the introduction to this thesis, screening is organized in an opportunistic fashion in the United States.[345] A large proportion of elderly people in the United States have never received any screening. With neither published data on the effect of screening for this population subgroup nor clear guidelines,[246] primary care practitioners were left to their own good judg-

ment in deciding what to recommend for these persons. Synthesizing knowledge of the natural history of colorectal cancer, screening test performance, and life expectancy according to comorbidity classes,[346] our study was the first to show that screening may be effective and cost-effective for previously unscreened people well beyond age 75 years.

Methodologically, the same strengths and limitations for the model apply to this study as above for other studies. We assumed no relationship between adenoma or cancer risk and comorbidity status. In elderly persons, this relationship is potentially more important for outcomes than for 50-75 year-old persons, because of higher other-cause mortality risk. Our assumption of independency may have overestimated screening benefits.

There were other limitations. First, the applied comorbidity classes were relatively broad in dividing persons into just three separate categories and factoring in only the most common medical conditions. Alternative co-morbidity scores may be more common in practice, such as the Charlson comorbidity index. There were also no empirical data available at the time of study conduct regarding the effect of screening in elderly unscreened persons. In recent years, one case-control study of 623 cases did suggest that screening may have very significant effects among elderly in general, but these findings need to be replicated by others.[347] Further, a blank screening history was the only risk factor considered for elderly persons, while there are many other known risk factors for colorectal cancer, including irregular screening exposure. Although the present study did not cover these alternative risk factors, the question to personalized stopping ages was addressed more comprehensively by the first author in a follow-up study.[348] Finally, we did not assess preferences and utilities from a personalized perspective, which would be methodologically more challenging.

A strength of our work on the benefit and cost-effectiveness of colonoscopy surveillance in patients with colorectal adenoma (Chapter 11), is that we validated our model to the best available data as identified by expert panels on adenoma risk during the initial surveillance period.[139] In itself, these data are difficult to interpret and do not provide clear-cut directions for optimal surveillance strategies. As mentioned above, previous studies have assessed the cost-effectiveness of adenoma surveillance using similar data sources, but none externally validated their model.[324,325,349] Our model matched well with the published adenoma detection rates during initial surveillance examination for most patients, except in detection of advanced adenomas for patients with high-risk adenomas at baseline. In sensitivity analyses, we evaluated alternative model variants to compensate for this suboptimal fit. Overall, there was a great consistency across evaluated models in suggesting low- and high-intensity surveillance to be both effective and cost-effective.

A limitation of the empirical work used to inform our study is the limited study sizes. Most studies reported small numbers of observed cancers. Therefore, carcinogenic risk for new-onset adenomas after adenoma removal is not clear. Our assumed conversion rates were based on standard assumptions regarding natural history, informed by large randomized controlled screening trials. As the base-case model validation suggested, the assumptions may not hold for patients with a history of high-risk adenomas.

A limitation of our study is that we only considered adenoma size and multiplicity to assess recurrence risk. We could not separately evaluate risk and recommendations for patients classified as high-risk for adenomas with high-grade dysplasia, (tubulo-)villous histology, or serrated features. Adenomas over 20mm in diameter could also not be studied separately, due to our model configuration. NCI pooling studies suggest that there may be a risk difference for high-risk patients with only tubular adenomas versus adenoma with alternative histology.[54] Evidence also suggests, however, that there is a strong correlation between size and presence of high-grade dysplasia, which may limit the loss of information by incorporating size alone.[53]

PRACTICAL IMPLICATIONS

Promote informed screening participation

The main conclusion from our work for the American Cancer Society can only be that the public health benefit of increased colorectal cancer screening, or screening in general, is potentially very substantial. We estimated that screening may reduce colorectal cancer-related deaths in the United States by 30-60% if screening uptake is increased to 80-100% (a 50-100% reduction in nonuse of screening). We believe these findings warrant further promotion of screening in the United States, and support the initiative to increase colorectal cancer screening rates to 80% by 2018. The US National Colorectal Cancer Roundtable media campaign has already quoted our findings.[350]

The implications stretch beyond the United States situation. Our results suggest that screening may have substantial benefits elsewhere, although an important distinction to make is that in most screening programs worldwide, stool-based tests are the prevailing screening modality, while in the United States, colonoscopy is the dominant screening test. As we showed in Chapter 6, the effects of stool-based testing may be lower compared to the effects of a completed colonoscopy given that long-term adherence with fecal tests is well below 100%.

There is a balance of benefits and harms in screening that we paid less attention to in our work presented in Chapters 3-5. We believe that to provide people with

sound information regarding benefits as well as harms is critically important. A common argument against screening is that it medicalizes healthy people. Inherently in screening there is balance of a small minority who benefits compared to a large majority in whom screening has no effect at all. Even in relatively effective colorectal cancer screening programs, to prevent one cancer-related death, at least 50 people may have to undergo colonoscopy examinations without any mortality benefit. In an era of patient-centered care, it is essential that people are made aware of this balance. Thus, initiatives such as the 80% by 2018 campaign should by no means forgo individuals' preferences and impede on people's autonomy by providing information that is one-sided.[351] Physicians in the United States who benefit from selling preventive services may have to be disincentivized. In other countries, with organized screening programs which take screening promotion from the hands of care providers, policy makers should be aware that high screening participation by poorly informed people does not necessarily define a successful program.

Implement quality monitoring and assurance programs

The main implication from our work in Part III of this thesis, is that quality assurance is essential for optimizing colorectal cancer screening outcomes. We estimated that up to 33% of screening benefits may be lost due to suboptimal adherence in a randomized clinical trial setting (in persons who provide consent to undergo screening), that long-term colorectal cancer outcomes may differ up to 60% between lower and higher colonoscopy quality settings (for an up to 32% difference in screening benefit compared to no screening), and that up to 10% of fecal testing benefits may be lost if follow-up of positive test results is delayed by 12 months. Although the relationship between these factors and participant outcomes should come as no surprise, the strength of the association has not been previously quantified. The magnitude of the effects support the American Gastroenterological Association guidelines which identified programmatic screening adherence and quality as the key program performance indicators.[114,352]

In practice, programs differ in terms of their performance. Adherence to screening is highly variable across different screening programs,[5] similar to quality of colonoscopy,[353] and diagnostic follow-up of positive results.[354] Some settings may be faced with low participant adherence while in other settings colonoscopy quality is falling short. Quality monitoring is therefore the first critical step to effective quality management. Once program performance is assessed, tailored interventions can be tested and implemented to target potential program deficiencies. For each of the above program aspects, there already exist interventions that have been evaluated in the literature, such as the use of navigators,[355,356] telephone reminders,[357] electronic reminders,[238] process flow maps,[343] and endoscopist training programs.[295]

Measure adenoma detection rates, but use with care

Our findings of a strong inverse association between ADR and long-term health outcomes support recent recommendations to use ADR as a primary intra-procedure quality indicator for colonoscopy.[131] The true novelty of our work was in suggesting an inverse relationship also with net costs of care. Total estimated costs of screening and treatment were lower for patients from physicians with high ADRs than for physicians with low ADRs. This implies that initiatives to measure and improve colonoscopy quality may be highly cost-effective. In the United States, this may also contribute to achieving the Institute for Healthcare Improvement's triple aim to improve the health of population, while reducing per-capita cost of care. For other countries, the cost implications could be an important side-aim for improving colonoscopy quality through appropriate quality assurance programs.

It is important to note that our work does not directly support the use of payment modifiers or any other incentives for increasing ADRs. It is not clear what the effect of such policies is. Although the evidence for the validity of ADR as a quality indicator is quite strong, ADR has previously mentioned limitations as a quality measure. The effect of ADR-related interventions on primary health outcomes (cancer, mortality) should ideally be established by other studies before these are widely applied.[295,358] Further, the use of complementary quality indicators like the number of detected adenomas per colonoscopy and withdrawal time should be considered for more rigorous quality assessment, to the extent that these are not already used. Finally, professional societies should consider setting minimum sample sizes for ADR assessment and/or recommending potential adjustments for case-mix differences.[359]

Colonoscopy is the most efficacious screening test, but may not be acceptable to all

Our findings from Part III single out colonoscopy as the most effective colorectal cancer screening test for people willing to undergo the test. In Chapter 8, high-quality colonoscopy was more efficacious than fecal testing for both incidence and mortality reductions. Even in settings with lower colonoscopy quality, a colonoscopy would likely outperform fecal testing due to the suboptimal long-term adherence (Chapter 6). As we demonstrated in Chapter 9, there may be additional factors affecting long-term outcomes of fecal testing. Thus, when exclusively considering the long-term effectiveness of each strategy and assuming that people are willing to undergo screening colonoscopy, this should be the preferred strategy for screening. This reflects the standpoint of the American College of Gastroenterology and United States Multi-Society Task Force guidelines.[102,115]

Literature suggests that many people may not be willing to undergo screening colonoscopy,[85,86] or may at least prefer non-invasive tests a priori over colonos-

copy.[360] Thus, offering colonoscopy as the sole option for screening, may have negative effects on overall population adherence. Conversely, adding FIT as an adjuvant test to colonoscopy may dramatically increase screening rates, as observed in California.[95] While some professional societies have suggested to offer persons an informed choice of practically all available tests for screening, the evidence for the success of such strategies is contentious.[351] The best way to exploit both the high potential effects of colonoscopy and the high acceptability of fecal tests may be to offer tests sequentially or in rank order from more to less effective, however, this is still to be investigated.

There are other criteria than effectiveness to consider when deciding which test strategy is most suitable for screening (**Chapter 1**). It will be important to evaluate the harms and cost-effectiveness of colonoscopy screening compared to sFOBT screening with actual adherence. There may also be practical factors that put constraints on the feasibility of offering colonoscopy for all, such as limited capacity or budgetary restrictions. In the Netherlands there currently is insufficient colonoscopy capacity for nation-wide colonoscopy screening. Past modeling studies have suggested that FIT screening is optimal in this case.[361] Finally, available evidence for a favorable benefits/harms ratio for colonoscopy may be valued differently by policy makers from various settings. While American expert groups have endorsed colonoscopy screening for years, European experts have been reluctant to take a similar standpoint in regard of the limited availability of experimental data on the benefits and harms of colonoscopy.[113] This reflects a principal standpoint on the quality of evidence, the debate of which is beyond the scope of this work.

Personalize stopping ages based on screening history and comorbidity status

The decision analysis on screening in elderly unscreened people addressed a very specific question and has practical implications. We found that screening may be considered for previously unscreened older persons depending on whether a person is in good health. Although United States Preventive Services guidelines have already suggested this in 2008,[362] the value of our work is in taking a model to classify people's overall health status and specifying corresponding screening cessation ages. Especially in settings where screening is organized in an opportunistic fashion the findings from our study are informative. Even in organized screening settings, however, policymakers could consider to offer screening beyond the usual stopping age of 75 years in people who have not been screened regularly before age 75 years.

More indirect, our work further suggests that a person's screening history is highly predictive of colorectal cancer risk and screening benefit. In the decision analysis for the United States Preventive Services Task Force,[259] our model partners showed

that screening persons beyond 75 years had no favorable benefit-to-harms ratio if persons were already screened for 25-30 years. However, we found that with no prior screening, the balance may be quite different. This suggests that future analyses and guidelines for more individualized screening should certainly consider past screening adherence as a predictor of screening benefit. Recent guideline recommendations also state this more explicitly.[103]

Surveillance should be based on adequately measured adenoma characteristics

Our work in Chapter 11 also has some implications for surveillance practice. Our results suggest that currently recommended surveillance in the United States for patients with low- and high-risk adenomas is effective and cost-effective compared to less intensive surveillance.[139] Previous suggestions that higher colonoscopy quality may lead to overuse of resources are not supported by our results, although intermediate-term results as assessed in the European Polyp Surveillance study may be different.[141] Whether more intensive surveillance for patients with very large polyps, or more than 5 small or 3 large adenomas is also cost-effective, as recommended in Europe, remains unclear. These strategies were not evaluated by us, and thus, cannot be dismissed as inappropriate based on our findings.

An indirect implication of our work is that assessment of adenoma size and multiplicity is requisite. In practice, polyps are often removed in piece-meal fashion, and many settings collect all detected polyps by organ section in a single jar for pathology review. This does not allow for accurate assessment of adenoma multiplicity. Size measurement is often foregone based on the argument that it is difficult to measure accurately. Opportunistically, however, inaccurate measurement is to be preferred over no measurement at all as the best of two imperfect options; even imprecise size estimates may allow for appropriate management of many more patients than no measurement at all.

FUTURE RESEARCH

Important research is being conducted to compare the benefits of the most common screening methods and surveillance strategies. Our suggestions for future research span several related subjects, including the cost-effectiveness of colonoscopy versus fecal testing with observed adherence, the impact of screening on life-years gained, development of more sensitive non-invasive screening tests, the effect of improved colonoscopy quality scores, and the potential for further personalization of screening and surveillance. Each of these research directions is briefly discussed below.

Cost-effectiveness of colonoscopy versus fecal colorectal cancer testing

To complement our work presented in Chapter 6, more research is needed on the benefit-to-harms ratio and cost-effectiveness of colonoscopy versus sFOBT with actual observed participant adherence rates. Follow-up studies could also look at the break-even point for adherence with colonoscopy screening at which this strategy would become equivalent to sFOBT screening in terms of health benefits. Exploratory analyses with the NCS data suggest that the break-even point for colonoscopy adherence may be between 50-60%. Finally, discrete choice experiments could be conducted to evaluate how offering screening tests in rank order from most to least effective, i.e. offering colonoscopy as the preferred test, and offering alternatives only if colonoscopy is rejected, influences people's behavior. Outcomes of this research could be combined with outcomes of effectiveness and cost-effectiveness research to estimate the ultimate impact of rank-ordered choice offers on primary screening outcomes.

All-cause mortality impact of screening

In future years, researchers should try to ascertain more accurate estimates of the all-cause mortality impact of colorectal cancer screening. The impact of screening on life-years gained has not been established to this date, while the assumed high impact of screening is one of the main drivers for its cost-effectiveness. Several new screening trials are underway in the United States and Europe.[86,121] Although none of these studies are powered on their own to find significant all-cause mortality effects, future meta-analyses should provide more clarity on the matter. If no effect on all-cause mortality is found, it may be illuminating to compare mortality rates for several competing causes of death for screened and unscreened subjects to identify how the benefit in terms of averted colorectal cancer deaths is offset. The ultimate implication of a lower than currently assumed all-cause mortality impact, is that cost-effectiveness models should downward adjust the assumed survival of patients with screen-detected adenomas and cancers, and possibly increase the assumed health care costs associated with the life-years gained.

Adenoma detection rate and other quality indicators

Future observational studies should assess whether improving ADRs improves patient's health. As we suggested in the last paragraph, stimulating higher ADRs may not actually improve quality, due to potential for gaming. The preferred study design for this may be a retrospective cohort study, because physicians would not be influenced by an awareness of being studied with respect to exam quality. If such a study would not find a substantial effect of ADR improvement on health outcomes, then follow-up research should be considered for developing composite quality

measures which integrate multiple distinct quality indicators, assessing minimum individual sample sizes for more precise quality measurement, and adjusting for case-mix characteristics. Pooled analyses with sufficient outcome points may be required to determine sample size thresholds for adequate quality measurement.

Test technology and genetics

Ongoing research is also needed for the development of more effective and specific non-invasive screening tests. In the last paragraph, we suggested that getting people to do colonoscopy may result in superior outcomes to annual fecal testing, due to difficulty in assuring persistent high adherence with the latter form of testing. For more sensitive and specific non-invasive forms of testing, this argument may not hold. These tests could reduce harms and the frequency of required testing, and thereby take away some of the current negative sentiments towards screening.

Breakthroughs in genetic research may be required for substantial forward leaps in test technology. While the genetic component of colorectal cancer is believed to be 12-35%,[363,364] most of the heritability is still unexplained.[365] Genome-wide association studies have discovered many genetic risk factors associated with colorectal cancer, but these explain only a fraction of the estimated heritability in colorectal cancer. Another promising area of research seems to be related to the human microbiome. Recent studies have found associations with colorectal cancer, but future studies should look closer at interactions with host factors.[48]

Replication of the impact of time to diagnostic testing

Our study on the effect of time to diagnostic follow-up of positive fecal tests suggested that there are clinically significant consequences of longer time to follow-up on colorectal cancer mortality. Empirical studies are needed to replicate these model predictions. Our work has already invoked a large-scale study by our research partners at Kaiser Permanente California.[366] Initial findings from the study as presented at a conference last year suggest that effect of colonoscopy timing is less linear than we assumed. However, the study confirmed our findings that delays of 6 months or more may have substantial effects on cancer incidence and advanced-stage cancer incidence.

Uncertainties in management of adenoma patients

The European Polyp Surveillance trial will provide essential evidence for the benefit of recommended surveillance in the mid-long term. This evidence will also be important for validation of long-term model projections. To further inform management of adenoma patients, future studies should aim to assess the potential for further personalization based on adenoma characteristics. Benefits of 1 year surveillance

intervals for patients with over 5 small adenomas, 3 or more large adenomas, and 1 very large adenoma are unclear, while already recommended by some European expert groups. Large observational studies may be needed to compare recurrent or residual adenoma findings for various surveillance intervals in these relatively small patient subgroups.

Individualized screening

In general, individualized screening is a central area for future study. Our work in Part IV looked at just two specific factors predictive of cancer risk (screening and adenoma history). In reality, there are many other risk factors for colorectal cancer and overall life expectancy. Future modeling studies should look at both start ages, stopping ages, and intervals for screening and surveillance, taking into consideration all significant risk factors, including but not limited to classical risk factors such as red meat consumption, alcohol, smoking and the use of aspirin. Information may be derived from quantitative stool-test results measuring (hemo)globin levels on a continuous scale, and, as we identified in this thesis, adenoma detection rates. With technological advancements there will likely also be increasing amounts of personal data including genetic and behavioral data available, at decreasing cost. Future research should assess the potential for both efficiency and effectiveness gains from more personalized screening strategies and guidelines. Initial studies by our group indicate that there likely is great potential.[348,367]

There are practical issues around personalized screening which deserve further attention in future research. First, personalized medicine is complex. Implementation of personalized guidelines may be problematic without technical tools to assist health care practitioners. Hence, if scientists and policy makers agree on the value of more personalized screening, researchers should also look for ways to implement them. A potential way for decision scientists to help effectuate personalized screening is to provide health care providers or screening organizations with a generic matrix containing starting ages, stopping ages and intervals for screening according to classes of relative disease risk and overall health status. Standardized risk tools could be used to allocate people on this matrix given specific personal risk characteristics. The matrix would then spell out the appropriate corresponding test strategy.

Second, there may be measurement problems related to personal risk assessment. Many risk factors are not independent. A person's risk may therefore only be estimable using regression models incorporating a multitude of factors, to control for dependency and interactions. However, with many included variables, there is a risk of statistical overfitting (number of parameters > number of observations). With real-time data collection via smart applications, people's risk status may change continuously. It may neither be feasible nor desirable to process all this information

in future screening models. Therefore, algorithms may have to be developed to select the most important person-level factors for informing individualized screening.

Finally, closer risk profiling or monitoring for prevention purposes may be too time-consuming for doctors and too intrusive for screening subjects. Before implementation should be considered, research is needed to estimate population adherence with personalized screening regimens. Researchers should evaluate how adherent people experience being faced with periodic health status updates and corresponding recommended screening strategies.

Future modeling

We want to underscore the anticipated persistent importance of modeling analyses in a changing landscape of screening. With expected shifts towards more individualized screening, potential improvements in cancer therapy (e.g. immunotherapy, cas9 gene-editing),[368,369] new available test technologies, likely changes in costs of screening and treatment, and ever tighter health care budgets, decision analyses will remain critical. Modeling should be informed by empirical data to the extent possible to improve the accuracy of predictions. Multi-disciplinary collaborations of modelers and health care practitioners remain vital to focus modeling on relevant questions for actual practice.

Important areas for improvement of the MISCAN-colon model include a stronger evidence-basis for the benefit of screening in terms of life-years gained, a stronger evidence basis for the effect of colonoscopy on proximal disease (in the right part of the colon), an update of the natural history model for other-than-traditional pathways to cancer and the molecular characteristics of lesions, and finally, as identified in Chapter 11, the potential inclusion of a person-specific factor for the rate of adenoma progression.

CONCLUSIONS

A remarkable achievement of the last century for public health has been the possibility to diagnose cancers in an early, more treatable stage. Colorectal cancer screening is potentially one the most effective secondary prevention methods, and ever since trials have unanimously proven the effectiveness of both fecal and invasive testing for the disease, colorectal cancer screening has been widely promoted. However, there are still notable uncertainties related to colorectal screening, some of which have been addressed in this thesis. Potential population-level effects of screening have not yet been assessed for many countries, the comparative programmatic performance of current stool-based tests versus colonoscopy are unknown, as are the

most effective ways to improve screening program performance, including the most powerful parameters for further personalization.

In Part II of this thesis, we estimated that screening has the potential to reduce colorectal cancer mortality in the United States by up to 60%. In Part III, we identified several important modifiable determinants of screening effectiveness in programmatic participant adherence, colonoscopy quality as reflected in adenoma detection rates, and time to follow-up of positive stool-based colorectal cancer test results. We also found, for persons prepared to undergo colonoscopy and fecal testing, that fecal testing with observed longitudinal adherence may be less effective than colonoscopy. Finally, in Part IV we found that screening may often be warranted beyond age 75 years in people who have not been previously screened, and that surveillance with 3-5 year intervals is cost-effective for adenoma patients.

In the general discussion of Part V, we made several recommendations. First, we expressed our support for promotion of screening as a highly effective public health intervention, on the condition that potential participants are informed properly on the balance of benefits and harms. We further recommended that quality assurance programs are implemented in screening practice, where the first important aim should be to measure a variety of relevant quality indicators, including screening adherence rates, time to follow-up in patients with positive stool test results, and ADRs. ADRs should be used with caution as the exclusive (intra-procedure) colonoscopy quality metric for quality improvement due to uncertainty regarding potential adverse effects. We further recommended that screening is considered beyond age 75 years if patients were not previously screened, and that 3-5 yearly colonoscopy surveillance is continued to be offered to patients with high- and low-risk adenomas, respectively. For implementation of effective surveillance policies, we recommended that physicians always attempt to assess adenoma size and multiplicity.

More research is still needed to assess whether offering colonoscopy as the preferred test for screening, and offering stool-based testing only if colonoscopy is denied, is an effective and cost-effective strategy to exploit both the high potential effects of colonoscopy and high acceptability of fecal testing. We further proposed as key areas for future research, the all-cause mortality effects of screening, the health impact of policies to stimulate higher ADRs, the search for novel DNA markers and improved non-invasive screening tests, and finally, the broad area of personalization in screening. The potential for effectiveness and efficiency gains through more personalized screening seems very substantial.

Many critical studies are ongoing at this point, that will shape the future of colorectal cancer screening. If advances in treatment will not already dramatically improve disease survival rates, developments in risk profiling may contribute to more effective, non-invasive strategies of screening that will further reduce the burden of

disease. We are optimistic that as the understanding of colorectal cancer causes and natural history further develops, the prospects for future colorectal cancer patients will also continue to improve.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65(2):87-108.
2. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med.* 1988;319(9):525-532.
3. Benson V, J Patnick, AK Davies, MR Nadel, RA Smith, WS Atkin. Colorectal cancer screening: A comparison of 35 initiatives in 17 countries. *International Journal of Cancer.* 2008;122:11.
4. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut.* 2015;64(10):1637-1649.
5. Klabunde C, Blom J, Bulliard JL, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. *J Med Screen.* 2015;22(3):119-126.
6. Rosen G. *A history of public health.* Baltimore, MD: The Johns Hopkins University Press; 1993.
7. Hippocrates. *Airs, Waters, Places.* Ann Arbor, MI: University of Michigan Library; 1981.
8. McNeill WH. *Plagues and peoples.* New York, NY: Knopf Doubleday Publishing Group; 2010.
9. Petty W. *Natural and political observations, mentioned in a following index, and made upon the bills of mortality by John Graunt.* 1662.
10. Morabia A, Zhang FF. History of medical screening: from concepts to action. *Postgrad Med J.* 2004;80(946):463-469.
11. Brown WJ. *Syphilis and other venereal disease.* Cambridge, MA: Harvard University Press; 1970.
12. Barringer T. The incidence of glycosuria and diabetes in New York City between 1902 and 1907. *Arch Intern Med.* 1909(3):295–298.
13. Striker C. *Famous faces in diabetes.* Boston, MA: G K Hall & Co; 1961.
14. Wilkerson HL, Krall LP. Diabetes in a New England town; a study of 3,516 persons in Oxford, Mass. *J Am Med Assoc.* 1947;135(4):209-216.
15. American Cancer Society. *Proceedings of the conference on cancer detection.* Portsmouth, NH1949.
16. Erickson C. Exfoliate cytology in mass screening for uterine cancer. *CA Cancer J Clin.* 1955(5):63-64.
17. Lockhart-Mummery J, Dukes C. The precancerous changes in the rectum and colon. *Surg Gynecol Obstet.* 1927;36:591–596.
18. Gilbertsen VA, Nelms JM. The prevention of invasive cancer of the rectum. *Cancer.* 1978;41(3):1137-1139.
19. Hertz RE, Deddish MR, Day E. Value of periodic examinations in detecting cancer of the rectum and colon. *Postgrad Med.* 1960;27:290-294.
20. Greegor DH. Diagnosis of large-bowel cancer in the asymptomatic patient. *JAMA.* 1967;201(12):943-945.
21. Wolff WI, Shinya H. Polypectomy via the fiberoptic colonoscope. Removal of neoplasms beyond reach of the sigmoidoscope. *N Engl J Med.* 1973;288(7):329-332.
22. 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):1541-1549.
23. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology.* 1997;112(2):594-642.
24. Von Karsa L, Anttila A, Ronco G, et al. *Cancer screening in the European Union; Report on the implementation of the Council Recommendation on cancer screening; First Report.* European Commission;2003.
25. de Koning HJ, Meza R, Plevritis SK, et al. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2014;160(5):311-320.

26. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2014;160(5):330-338.
27. Porter D. *The History of Public Health and the Modern State.* Atlanta, GA: Rodopi; 1994.
28. 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(20):1412-1422.
29. Taksler GB, Keshner M, Fagerlin A, Hajizadeh N, Braithwaite RS. Personalized estimates of benefit from preventive care guidelines: a proof of concept. *Ann Intern Med.* 2013;159(3):161-168.
30. Commission on Chronic Illness. *Chronic Illnesses in the United States.* Vol I. Cambridge, MA: Harvard University Press; 1957.
31. Wilson JMG, Jungner G. Principles and practice of screning for disease. Geneva, CH: World Health Organization; 1968: http://whqlibdoc.who.int/php/who_php_34.pdf.
32. Public Health Resource Unit. *Guidance for public health and commissioners.* Sheffield, UK: NHS Cancer Screening Programmes;2008.
33. de Visser M, van Ballegooijen M, Bloemers SM, et al. Report on the Dutch consensus development meeting for implementation and further development of population screening for colorectal cancer based on FOBT. *Cell Oncol.* 2005;27(1):17-29.
34. Andermann A, Blancquaert I Fau - Beauchamp S, Beauchamp S Fau - Dery V, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. (1564-0604 (Electronic)).
35. de Koning HJ. *The mysterious mass(es)* 2009.
36. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive services task force. *Epidemiol Rev.* 2011;33:20-35.
37. World Health Organization. *Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis.* Geneva: World Health Orginzation; 2003.
38. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *Jama.* 1996;276(15):1253-1258.
39. National Institute for Health and Care Excellence. *Developing NICE guidelines: the manual.* National Institute for Health and Care Excellence,;2014.
40. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics.* 2008;26(9):733-744.
41. College voor Zorgverzekeringen. *Kosteneffectiviteit in de zorg.* 2013.
42. World Health Organization. Global Burden Of Disease; Global Health Estimates. 2014; http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. Accessed February 3, 2015.
43. Peto R, Darby S Fau - Deo H, Deo H Fau - Silcocks P, Silcocks P Fau - Whitley E, Whitley E Fau - Doll R, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. (0959-8138 (Print)).
44. International Agency for Research on Cancer (WHO). GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence in 2012. 2012; http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed February 2, 2015.
45. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol Biomarkers Prev.* 2010;19(8):1893-1907.

46. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol.* 2001;96(10):2992-3003.
47. Aaltonen L, Johns L, Jarvinen H, Mecklin JP, Houlston R. Explaining the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. *Clin Cancer Res.* 2007;13(1):356-361.
48. Kostic AD, Gevers D, Pedamallu CS, et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res.* 2012;22(2):292-298.
49. Colditz GA, Atwood KA, Emmons K, et al. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. Risk Index Working Group, Harvard Center for Cancer Prevention. *Cancer Causes Control.* 2000;11(6):477-488.
50. Kuntz KM, Lansdorp-Vogelaar I, Rutter CM, et al. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. *Med Decis Making.* 2011;31(4):530-539.
51. Bosman FT, Carneiro F, Hruban RH, Theise ND. *WHO Classification of Tumours of the Digestive System.* Geneva: WHO Press; 2010.
52. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. *Cancer.* 1975;36(6):2251-2270.
53. Aldridge AJ, Simson JN. Histological assessment of colorectal adenomas by size. Are polyps less than 10 mm in size clinically important? *Eur J Surg.* 2001;167(10):777-781.
54. Martinez ME, Baron JA, Lieberman DA, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology.* 2009;136(3):832-841.
55. Martinez ME, Thompson P, Messer K, et al. One-year risk for advanced colorectal neoplasia: U.S. versus U.K. risk-stratification guidelines. *Ann Intern Med.* 2012;157(12):856-864.
56. Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. *Am J Gastroenterol.* 2012;107(9):1315-1329; quiz 1314, 1330.
57. Lash RH, Genta Rm Fau - Schuler CM, Schuler CM. Sessile serrated adenomas: prevalence of dysplasia and carcinoma in 2139 patients. (1472-4146 (Electronic)).
58. Rex DK, Ahnen Dj Fau - Baron JA, Baron Ja Fau - Batts KP, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. (1572-0241 (Electronic)).
59. 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.
60. Cairns SR, Scholefield JH, Steele RJ, et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). *Gut.* 2010;59(5):666-689.
61. van Heijningen EM, Lansdorp-Vogelaar I, Steyerberg EW, et al. Adherence to surveillance guidelines after removal of colorectal adenomas: a large, community-based study. *Gut.* 2015;64(10):1584-1592.
62. Sheffield KM, Han Y, Kuo YF, Riall TS, Goodwin JS. Potentially inappropriate screening colonoscopy in Medicare patients: variation by physician and geographic region. *JAMA Intern Med.* 2013;173(7):542-550.
63. Goodwin JS, Singh A, Reddy N, Riall TS, Kuo YF. Overuse of screening colonoscopy in the Medicare population. *Arch Intern Med.* 2011;171(15):1335-1343.
64. Yabroff KR, Klabunde CN, Yuan G, et al. Are physicians' recommendations for colorectal cancer screening guideline-consistent? *J Gen Intern Med.* 2011;26(2):177-184.
65. Grubbe E. Priority in therapeutic use of X-rays. *Radiology.* 1933;21:156-162.

66. Farber S, Diamond LK. Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid. *N Engl J Med.* 1948;238(23):787-793.
67. NCCN guidelines v2.2016 panel. *Colon/Rectal Cancer* 2016.
68. Yabroff KR, Mariotto Ab Fau - Feuer E, Feuer E Fau - Brown ML, Brown ML. Projections of the costs associated with colorectal cancer care in the United States, 2000-2020. (1057-9230 (Print)).
69. Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey. 2012. Accessed March 3, 2015.
70. Schoen C, Osborn R, Squires D, Doty MM. Access, affordability, and insurance complexity are often worse in the United States compared to ten other countries. *Health Aff (Millwood).* 2013;32(12):2205-2215.
71. The World Bank. Health expenditure, total (% of GDP) 2014; http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS. Accessed March 3, 2015.
72. Dieleman JL, Graves C, Johnson E, et al. Sources and Focus of Health Development Assistance, 1990-2014. *Jama.* 2015;313(23):2359-2368.
73. Sung JJ, Lau JY, Goh KL, Leung WK, Asia Pacific Working Group on Colorectal C. Increasing incidence of colorectal cancer in Asia: implications for screening. *Lancet Oncol.* 2005;6(11):871-876.
74. Sung JJ, Ng SC, Chan FK, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. *Gut.* 2015;64(1):121-132.
75. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2013 Sub (1973-2011) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2012 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2014, based on the November 2013 submission.
76. Douillard JY, Oliner KS, Siena S, et al. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. *N Engl J Med.* 2013;369(11):1023-1034.
77. Saltz LB, Clarke S, Diaz-Rubio E, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol.* 2008;26(12):2013-2019.
78. Tol J, Koopman M, Cats A, et al. Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. *N Engl J Med.* 2009;360(6):563-572.
79. Van Cutsem E, Lenz HJ, Kohne CH, et al. Fluorouracil, leucovorin, and irinotecan plus cetuximab treatment and RAS mutations in colorectal cancer. *J Clin Oncol.* 2015;33(7):692-700.
80. Dickinson W. *Inspection of the rectum and sigmoid flexure by mechanical means.* . 1901.
81. Overholt BF. Clinical experience with the fibersigmoidoscope. *Gastrointest Endosc.* 1968;15(1):27.
82. 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.
83. Gupta S, Halm EA, Rockey DC, et al. Comparative effectiveness of fecal immunochemical test outreach, colonoscopy outreach, and usual care for boosting colorectal cancer screening among the underserved: a randomized clinical trial. *JAMA Intern Med.* 2013;173(18):1725-1732.
84. 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.
85. 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):575-582.

86. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. *N Engl J Med.* 2012;366(8):697-706.
87. 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):347-357.
88. 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.
89. Shapiro JA, Klabunde CN, Thompson TD, Nadel MR, Seeff LC, White A. Patterns of colorectal cancer test use, including CT colonography, in the 2010 National Health Interview Survey. *Cancer Epidemiol Biomarkers Prev.* 2012;21(6):895-904.
90. 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.
91. Jensen CD, Corley DA, Quinn VP, et al. Fecal Immunochemical Test Program Performance Over 4 Rounds of Annual Screening: A Retrospective Cohort Study. *Ann Intern Med.* 2016;164(7):456-463.
92. Crotta S, Segnan N, Paganin S, Dagnes B, Rosset R, Senore C. High rate of advanced adenoma detection in 4 rounds of colorectal cancer screening with the fecal immunochemical test. *Clin Gastroenterol Hepatol.* 2012;10(6):633-638.
93. Parikh RB, Prasad V. Blood-Based Screening for Colon Cancer: A Disruptive Innovation or Simply a Disruption? *JAMA.* 2016;315(23):2519-2520.
94. Klabunde CN, Joseph DA, King JB, White A, Plescia M. *Vital Signs: Colorectal Cancer Screening Test Use — United States, 2012.* 2013.
95. Lee JK, Levin TR, Corley DA. The road ahead: what if gastroenterologists were accountable for preventing colorectal cancer? *Clin Gastroenterol Hepatol.* 2013;11(3):204-207.
96. Mehta SJ, Jensen CD, Quinn VP, et al. Race/Ethnicity and Adoption of a Population Health Management Approach to Colorectal Cancer Screening in a Community-Based Healthcare System. *J Gen Intern Med.* 2016.
97. Pox CP, Altenhofen L, Brenner H, Theilmeier A, Von Stillfried D, Schmiegel W. Efficacy of a nationwide screening colonoscopy program for colorectal cancer. *Gastroenterology.* 2012;142(7):1460-1467 e1462.
98. National Institute for Public Health and Environment. Bevolkingsonderzoek Darmkanker. 2015; http://www.rivm.nl/dsresource?objectid=rivmp:282150&type=org&disposition=inline&ns_nc=1. Accessed July 27, 2016.
99. Doubeni CA, Corley DA, Zauber AG. Colorectal Cancer Health Disparities and the Role of US Law and Health Policy. *Gastroenterology.* 2016;150(5):1052-1055.
100. Subramanian S, Tangka FK, Hoover S, et al. Costs of planning and implementing the CDC's Colorectal Cancer Screening Demonstration Program. *Cancer.* 2013;119 Suppl 15:2855-2862.
101. Johnson DA, Barclay RL, Mergener K, et al. Plasma Septin9 versus fecal immunochemical testing for colorectal cancer screening: a prospective multicenter study. *PLoS One.* 2014;9(6):e98238.
102. Levin B, DA Lieberman, SJ Winawer et al. Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline From the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology.* 2008;134:26.

103. U. S. Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2016;315(23):2564-2575.
104. 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):2576-2594.
105. Chiu HM, Chen SL, Yen AM, et al. Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. *Cancer.* 2015;121(18):3221-3229.
106. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. *Gut.* 2015;64(5):784-790.
107. Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ.* 2014;348:g2467.
108. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. *Epidemiol Rev.* 2011;33:88-100.
109. Armaroli P, Villain P, Suonio E, et al. European Code against Cancer, 4th Edition: Cancer screening. *Cancer Epidemiol.* 2015;39 Suppl 1:S139-152.
110. Wilschut JA, Hol L, Dekker E, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. *Gastroenterology.* 2011;141(5):1648-1655 e1641.
111. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, Wilschut JA, Zauber AG, van Ballegooijen M. Stool DNA testing to screen for colorectal cancer in the Medicare population: a cost-effectiveness analysis. *Ann Intern Med.* 2010;153(6):368-377.
112. Council of the European Union. Council recommendations of 2 December 2003 on cancer screening. *Official Journal of the European Union.* 2003;L327:34-38.
113. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, 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):51-59.
114. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *CA Cancer J Clin.* 2008;58(3):130-160.
115. Rex DK, Johnson DA, Anderson JC, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. *Am J Gastroenterol.* 2009;104(3):739-750.
116. van Hees F, Zauber AG, van Veldhuizen H, et al. The value of models in informing resource allocation in colorectal cancer screening: the case of The Netherlands. *Gut.* 2015;64(12):1985-1997.
117. 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.
118. Holme O, Loberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. *JAMA.* 2014;312(6):606-615.
119. Schoen RE, Pinsky PF, Weissfeld JL, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med.* 2012;366(25):2345-2357.

120. Segnan N, Armaroli P, Bonelli L, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. *J Natl Cancer Inst.* 2011;103(17):1310-1322.
121. Bretthauer M, Kaminski MF, Loberg M, et al. Population-Based Colonoscopy Screening for Colorectal Cancer: A Randomized Clinical Trial. *JAMA Intern Med.* 2016;176(7):894-902.
122. ClinicalTrials.gov. Colonoscopy Versus Fecal Immunochemical Test in Reducing Mortality From Colorectal Cancer (CONFIRM). 2015; https://clinicaltrials.gov/ct2/show/NCT01239082. Accessed August 1, 2015.
123. Sonnenberg A, Delco F, Inadomi JM. Cost-effectiveness of colonoscopy in screening for colorectal cancer. *Ann Intern Med.* 2000;133(8):573-584.
124. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med.* 2014;370(14):1298-1306.
125. Kaminski MF, J. Regula, E. Butruk, et al. Quality Indicators for Colonoscopy and the Risk of Interval Cancer. *New England Journal of Medicine.* 2010;362(19):9.
126. Clark BT, Rustagi T, Laine L. What Level of Bowel Prep Quality Requires Early Repeat Colonoscopy: Systematic Review and Meta-Analysis of the Impact of Preparation Quality on Adenoma Detection Rate. *Am J Gastroenterol.* 2014;109(11):1714-1723.
127. Baxter NN, Sutradhar R, Forbes SS, Paszat LF, Saskin R, Rabeneck L. Analysis of administrative data finds endoscopist quality measures associated with postcolonoscopy colorectal cancer. *Gastroenterology.* 2011;140(1):65-72.
128. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *N Engl J Med.* 2006;355(24):2533-2541.
129. Shaukat A, Rector TS, Church TR, et al. Longer Withdrawal Time Is Associated With a Reduced Incidence of Interval Cancer After Screening Colonoscopy. *Gastroenterology.* 2015.
130. Hilsden RJ, Dube C, Heitman SJ, Bridges R, McGregor SE, Rostom A. The association of colonoscopy quality indicators with the detection of screen-relevant lesions, adverse events, and postcolonoscopy cancers in an asymptomatic Canadian colorectal cancer screening population. *Gastrointest Endosc.* 2015;82(5):887-894.
131. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. *Am J Gastroenterol.* 2015;110(1):72-90.
132. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. *Am J Gastroenterol.* 2006;101(4):873-885.
133. Imperiale TF, Monahan PO, Stump TE, Glowinski EA, Ransohoff DF. Derivation and Validation of a Scoring System to Stratify Risk for Advanced Colorectal Neoplasia in Asymptomatic Adults: A Cross-sectional Study. *Ann Intern Med.* 2015;163(5):339-346.
134. Shaukat A, Church TR, Shanley R, et al. Development and validation of a clinical score for predicting risk of adenoma at screening colonoscopy. *Cancer Epidemiol Biomarkers Prev.* 2015;24(6):913-920.
135. Yeoh KG, Ho KY, Chiu HM, et al. The Asia-Pacific Colorectal Screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects. *Gut.* 2011;60(9):1236-1241.
136. Chiu HM, Ching JY, Wu KC, et al. A Risk-Scoring System Combined With a Fecal Immunochemical Test Is Effective in Screening High-Risk Subjects for Early Colonoscopy to Detect Advanced Colorectal Neoplasms. *Gastroenterology.* 2016;150(3):617-625 e613.

137. Chiu SY, Chuang SL, Chen SL, et al. Faecal haemoglobin concentration influences risk prediction of interval cancers resulting from inadequate colonoscopy quality: analysis of the Taiwanese Nationwide Colorectal Cancer Screening Program. *Gut*. 2015.
138. Saini SD, van Hees F, Vijan S. Smarter screening for cancer: possibilities and challenges of personalization. *JAMA*. 2014;312(21):2211-2212.
139. 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(3):844-857.
140. Winawer SJ AZ, E Stuart et al. The prevention of colorectal cancer by colonoscopic polypectomy. *New England Journal of Medicine*. 1993;329(27):5.
141. Austin GL, Fennimore B, Ahnen DJ. Can colonoscopy remain cost-effective for colorectal cancer screening? The impact of practice patterns and the Will Rogers phenomenon on costs. *Am J Gastroenterol*. 2013;108(3):296-301.
142. Gupta S, Jacobs ET, Baron JA, et al. Risk stratification of individuals with low-risk colorectal adenomas using clinical characteristics: a pooled analysis. *Gut*. 2015.
143. Hennink SD, van der Meulen-de Jong AE, Wolterbeek R, et al. Randomized Comparison of Surveillance Intervals in Familial Colorectal Cancer. *J Clin Oncol*. 2015;33(35):4188-4193.
144. Tran AH, Man Ngor EW, Wu BU. Surveillance colonoscopy in elderly patients: a retrospective cohort study. *JAMA Intern Med*. 2014;174(10):1675-1682.
145. Loberg M, Kalager M, Holme O, Hoff G, Adami HO, Bretthauer M. Long-term colorectal-cancer mortality after adenoma removal. *N Engl J Med*. 2014;371(9):799-807.
146. Jover R, Bretthauer M, Dekker E, et al. Rationale and design of the European Polyp Surveillance (EPoS) trials. *Endoscopy*. 2016.
147. National Cancer Institute. Population-based Research Optimizing Screening through Personalized Regimens. 2011; http://appliedresearch.cancer.gov/networks/prospr/.
148. Cancer Intervention and Surveillance Modeling Network (CISNET). Colorectal Cancer Model Profiles. http://cisnet.cancer.gov/colorectal/profiles.html/. Accessed January 1, 2013.
149. American Joint Committee on Cancer. AJCC Cancer Staging Manual. 5th ed. Philadelphia: Lippincott - Raven Publishers; 1997: http://cancerstaging.org/references-tools/deskreferences/Documents/AJCC5thEdCancerStagingManual.pdf.
150. Sobin LH, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours*. 7 ed: Wiley-Blackwell; 2009.
151. Dukes C.E. The classification of cancer of the rectum. *Journal of Pathological Bacteriology* 1932;35(323).
152. Zauber AG, Knudsen AB, Rutter CM, et al. Cost-Effectiveness of CT Colonography to Screen for Colorectal Cancer. 2007. http://www.cms.gov/Medicare/Coverage/DeterminationProcess/downloads/id52TA.pdf.
153. Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63(2):317-325.
154. 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.
155. Lindholm E, Brevinge H, Haglind E. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. *Br J Surg*. 2008;95(8):1029-1036.
156. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med*. 2000;343(22):1603-1607.

157. 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.
158. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. *Gut.* 2012;61(7):1036-1040.
159. Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA, Ross ME, Colorectal Cancer Study G. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *N Engl J Med.* 2004;351(26):2704-2714.
160. 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.
161. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med.* 2014;370(14):1287-1297.
162. Hoff G, Grotmol T, Skovlund E, Bretthauer M, Norwegian Colorectal Cancer Prevention Study G. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. *Bmj.* 2009;338:b1846.
163. 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.
164. Kahi CJ, Imperiale TF, Juliar BE, Rex DK. Effect of screening colonoscopy on colorectal cancer incidence and mortality. *Clin Gastroenterol Hepatol.* 2009;7(7):770-775; quiz 711.
165. Doubeni CA, Weinmann S, Adams K, et al. Screening colonoscopy and risk for incident late-stage colorectal cancer diagnosis in average-risk adults: a nested case-control study. *Ann Intern Med.* 2013;158(5 Pt 1):312-320.
166. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med.* 2013;369(12):1095-1105.
167. Brenner H, Chang-Claude J, Jansen L, Knebel P, Stock C, Hoffmeister M. Reduced risk of colorectal cancer up to 10 years after screening, surveillance, or diagnostic colonoscopy. *Gastroenterology.* 2014;146(3):709-717.
168. Cotterchio M, Manno M, Klar N, McLaughlin J, Gallinger S. Colorectal screening is associated with reduced colorectal cancer risk: a case-control study within the population-based Ontario Familial Colorectal Cancer Registry. *Cancer Causes Control.* 2005;16(7):865-875.
169. Manser CN, Bachmann LM, Brunner J, Hunold F, Bauerfeind P, Marbet UA. Colonoscopy screening markedly reduces the occurrence of colon carcinomas and carcinoma-related death: a closed cohort study. *Gastrointest Endosc.* 2012;76(1):110-117.
170. Rex DK, Adler SN, Aisenberg J, et al. Accuracy of capsule colonoscopy in detecting colorectal polyps in a screening population. *Gastroenterology.* 2015;148(5):948-957 e942.
171. Winawer SJ, Stewart ET, Zauber AG, et al. A comparison of colonoscopy and double-contrast barium enema for surveillance after polypectomy. National Polyp Study Work Group. *N Engl J Med.* 2000;342(24):1766-1772.
172. Church TR, Ederer F, Mandel JS. Fecal occult blood screening in the Minnesota study: sensitivity of the screening test. *J Natl Cancer Inst.* 1997;89(19):1440-1448.
173. International Agency for Research on Cancer (WHO). Globocan 2012: Estimated Cancer Incidence, Mortality, and Prevalence Worldwide in 2012. 2012; http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx?cancer=colorectal. Accessed July 27, 2016.
174. National Center for Health Statistics. Historical Data, 1900-1998. 2009; http://www.cdc.gov/nchs/nvss/mortality_historical_data.htm. Accessed July 27, 2016.

175. Surveillance Epidemiology and End Results Program. SEER*Stat Software, version 5.3.1. Surveillance Research Program. 2014; http://seer.cancer.gov/seerstat/. Accessed January 1, 2015.
176. Centers for Disease Control and Prevention. Life tables. 2010; http://www.cdc.gov/nchs/products/life_tables.htm. Accessed October 28, 2013.
177. Rutter CM, Johnson EA, Feuer EJ, Knudsen AB, Kuntz KM, Schrag D. Secular trends in colon and rectal cancer relative survival. *J Natl Cancer Inst.* 2013;105(23):1806-1813.
178. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut.* 1982;23(10):835-842.
179. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. *Cancer.* 1982;49(4):819-825.
180. 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.
181. 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.
182. 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.
183. Blatt LJ. Polyps of the colon and rectum: incidence and distribution. *Dis Colon Rectum.* 1961;4:277-282.
184. 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.
185. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. *Cancer.* 1988;61(7):1472-1476.
186. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. *Ann Surg.* 1963;157:223-226.
187. 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.
188. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet.* 1996;348(9040):1472-1477.
189. Jorgensen OD, Kronborg O, Fenger C. A randomised study of screening for colorectal cancer using faecal occult blood testing: results after 13 years and seven biennial screening rounds. *Gut.* 2002;50(1):29-32.
190. 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.
191. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer.* 2004;111(4):633-639.
192. Rutter CM, Knudsen AB, Marsh TL, et al. Validation of Models Used to Inform Colorectal Cancer Screening Guidelines: Accuracy and Implications. *Med Decis Making.* 2016;36(5):604-614.
193. Zauber A, SJ Winawer, JD Waye et al. Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths. *New England Journal of Medicine.* 2012;366(8):10.
194. Holme O, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane Database Syst Rev.* 2013;9:CD009259.
195. van Ballegooijen M, Rutter CM, Knudsen AB, et al. Clarifying differences in natural history between models of screening: the case of colorectal cancer. *Med Decis Making.* 2011;31(4):540-549.

196. National Cancer Institute. SEER Cancer Statistics Factsheets: Colon and Rectum Cancer. 2014; http://seer.cancer.gov/statfacts/html/colorect.html/. Accessed September, 2014.
197. Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin.* 2014;64(2):104-117.
198. Edwards BK, Ward E, Kohler BA, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer.* 2010;116(3):544-573.
199. Yang DX, Gross CP, Soulos PR, Yu JB. Estimating the magnitude of colorectal cancers prevented during the era of screening: 1976 to 2009. *Cancer.* 2014.
200. Centers for Disease Control and Prevention. National Health Interview Survey. http://www.cdc.gov/nchs/nhis.htm. Accessed March, 2014.
201. Stock C, Knudsen AB, Lansdorp-Vogelaar I, Haug U, Brenner H. Colorectal cancer mortality prevented by use and attributable to nonuse of colonoscopy. *Gastrointest Endosc.* 2011;73(3):435-443 e435.
202. Meester RGS, Doubeni CA, Lansdorp-Vogelaar I, et al. Colorectal Cancer Deaths Attributable to Nonuse of Screening in the United States. *Annals of Epidemiology.* 2015;25(3):208-213.
203. American Cancer Society. The National Colorectal Cancer Roundtable. 2014; http://www.cancer.org/healthy/informationforhealthcareprofessionals/colonmdclinicansinformationsource/nationalcolorectalcancerroundtable/national-colorectal-cancer-roundtable. Accessed January, 2015.
204. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2008;149(9):659-669.
205. U.S. Census Bureau. National Population Projections. 2000; http://www.census.gov/population/projections/. Accessed January, 2015.
206. Berkeley Mortality Database. Lifetables by year of birth 1900-2000. http://www.demog.berkeley.edu/~bmd/states.html. Accessed January, 2015.
207. 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.
208. 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 Dis.* 2004;6(3):158-161.
209. Meester RGS, Doubeni CA, Lansdorp-Vogelaar I, et al. Colorectal Cancer Deaths Attributable to Nonuse of Screening in the United States. *Annals of Epidemiology.* 2015;In press.
210. Healthy People 2020. Disparities. http://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities. Accessed January, 2015.
211. Ladabaum U, Song K. Projected national impact of colorectal cancer screening on clinical and economic outcomes and health services demand. *Gastroenterology.* 2005;129(4):1151-1162.
212. Vogelaar I, van Ballegooijen M, Schrag D, et al. How much can current interventions reduce colorectal cancer mortality in the U.S.? Mortality projections for scenarios of risk-factor modification, screening, and treatment. *Cancer.* 2006;107(7):1624-1633.
213. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, van Ballegooijen M, Zauber AG, Jemal A. Contribution of screening and survival differences to racial disparities in colorectal cancer rates. *Cancer Epidemiol Biomarkers Prev.* 2012;21(5):728-736.

214. van Hees F, Habbema JD, Meester RG, Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG. Should colorectal cancer screening be considered in elderly persons without previous screening?: a cost-effectiveness analysis. *Ann Intern Med.* 2014;160(11):750-759.
215. Dinh T, Ladabaum U, Alperin P, Caldwell C, Smith R, Levin TR. Health benefits and cost-effectiveness of a hybrid screening strategy for colorectal cancer. *Clin Gastroenterol Hepatol.* 2013;11(9):1158-1166.
216. Wender R. Reaching 80% by 2018: What It's Really Going To Take. *2014 C5 Summit Presentations* [Powerpoint presentation]. 2014; http://c5nyc.org/2014-c5-summit-presentations/#sthash.2hJy6pdA.dpuf. Accessed January, 2015.
217. National Center for Health Statistics. Dat file documentation, National Health Interview Survey, 2010 (machine readable data file and documentation). Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention; 2014.
218. Meester RG, Doubeni CA, Zauber AG, et al. Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018. *Cancer.* 2015.
219. Fedewa SA, Ma J, Sauer AG, et al. How many individuals will need to be screened to increase colorectal cancer screening prevalence to 80% by 2018? *Cancer.* 2015;121(23):4258-4265.
220. Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *CA Cancer J Clin.* 2006;56(3):143-159; quiz 184-145.
221. Seeff LC, Richards TB, Shapiro JA, et al. How many endoscopies are performed for colorectal cancer screening? Results from CDC's survey of endoscopic capacity. *Gastroenterology.* 2004;127(6):1670-1677.
222. Klabunde CN, Cronin KA, Breen N, Waldron WR, Ambs AH, Nadel MR. Trends in colorectal cancer test use among vulnerable populations in the United States. *Cancer Epidemiol Biomarkers Prev.* 2011;20(8):1611-1621.
223. Rim SH, Joseph DA, Steele CB, et al. Colorectal cancer screening - United States, 2002, 2004, 2006, and 2008. *MMWR Suppl.* 2011;60(1):42-46.
224. Dorn SD, Wei D, Farley JF, et al. Impact of the 2008-2009 economic recession on screening colonoscopy utilization among the insured. *Clin Gastroenterol Hepatol.* 2012;10(3):278-284.
225. Koenig L, Gu Q. Growth of ambulatory surgical centers, surgery volume, and savings to medicare. *Am J Gastroenterol.* 2013;108(1):10-15.
226. Centers for Disease Control and Prevention. *Behavioral Risk Factor Surveillance System Survey Data.* In: Services USDoHaH, ed. Atlanta, GA2012.
227. Lieberman DA, Holub J, Eisen G, Kraemer D, Morris CD. Utilization of colonoscopy in the United States: results from a national consortium. *Gastrointest Endosc.* 2005;62(6):875-883.
228. Lieberman DA, De Garmo PL, Fleischer DE, Eisen GM, Helfand M. Patterns of endoscopy use in the United States. *Gastroenterology.* 2000;118(3):619-624.
229. Cooper GS, Kou Td Fau - Barnholtz Sloan JS, Barnholtz Sloan Js Fau - Koroukian SM, Koroukian Sm Fau - Schluchter MD, Schluchter MD. Use of colonoscopy for polyp surveillance in Medicare beneficiaries. (1097-0142 (Electronic)).
230. Laiyemo AO, Pinsky PF, Marcus PM, et al. Utilization and yield of surveillance colonoscopy in the continued follow-up study of the polyp prevention trial. *Clin Gastroenterol Hepatol.* 2009;7(5):562-567; quiz 497.
231. Boolchand V, Olds G, Singh J, Singh P, Chak A, Cooper GS. Colorectal screening after polypectomy: a national survey study of primary care physicians. *Ann Intern Med.* 2006;145(9):654-659.

232. Schoen RE, Pinsky PF, Weissfeld JL, et al. Utilization of surveillance colonoscopy in community practice. *Gastroenterology.* 2010;138(1):73-81.
233. Mysliwiec PA, Brown ML, Klabunde CN, Ransohoff DF. Are physicians doing too much colonoscopy? A national survey of colorectal surveillance after polypectomy. *Ann Intern Med.* 2004;141(4):264-271.
234. Centers for Disease Control and Prevention. Colorectal cancer screening - United States, 2002, 2004, 2006, and 2008. . *Morbidity and Mortality Weekly Report (MMWR).* 2011;60:42-46.
235. Holden DJ, Jonas DE, Porterfield DS, Reuland D, Harris R. Systematic review: enhancing the use and quality of colorectal cancer screening. *Ann Intern Med.* 2010;152(10):668-676.
236. Seeff LC, Royalty J, Helsel WE, et al. Clinical outcomes from the CDC's Colorectal Cancer Screening Demonstration Program. *Cancer.* 2013;119 Suppl 15:2820-2833.
237. Nadel MR, Royalty J, Shapiro JA, et al. Assessing screening quality in the CDC's Colorectal Cancer Screening Demonstration Program. *Cancer.* 2013;119 Suppl 15:2834-2841.
238. Humphrey LL, Shannon J, Partin MR, O'Malley J, Chen Z, Helfand M. Improving the follow-up of positive hemoccult screening tests: an electronic intervention. *J Gen Intern Med.* 2011;26(7):691-697.
239. Levin TR, Jamieson L, Burley DA, Reyes J, Oehrli M, Caldwell C. Organized colorectal cancer screening in integrated health care systems. *Epidemiol Rev.* 2011;33(1):101-110.
240. Centres for Disease Control and Prevention. March Is National Colorectal Cancer Awareness Month. 2014; http://www.cdc.gov/cancer/dcpc/resources/features/colorectalawareness/. Accessed May, 30, 2014.
241. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum.* 1953;9(3):531-541.
242. Hilsden RJ. Quantifying the benefit of screening colonoscopy. *Gastrointest Endosc.* 2011;73(3):444-446.
243. Muller AD, Sonnenberg A. Protection by endoscopy against death from colorectal cancer. A case-control study among veterans. *Arch Intern Med.* 1995;155(16):1741-1748.
244. Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. Association of colonoscopy and death from colorectal cancer. *Ann Intern Med.* 2009;150(1):1-8.
245. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health.* 1998;88(1):15-19.
246. U. S. Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2008;149(9):627-637.
247. Atkin W, R Edwards, J Cuzick. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet.* 2010;Online.
248. National Cancer Institute. Age-Specific U.S. Death Rates, 2006-2010. 2013; http://seer.cancer.gov/archive/csr/1975_2010/browse_csr.php. Accessed April 23,, 2014.
249. U.S. Census Bureau. Population Estimates by age, sex, race and Hispanic origin. 2012; http://www.census.gov/popest/data/index.html.
250. American Cancer Society. Colorectal cancer screening - state and federal coverage laws. 2013; http://www.cancer.org/cancer/colonandrectumcancer/moreinformation/colonandrectumcancerearlydetection/colorectal-cancer-early-detection-screening-coverage-laws. Accessed May, 30, 2014.
251. Rauscher GH, Johnson TP, Cho YI, Walk JA. Accuracy of self-reported cancer-screening histories: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2008;17(4):748-757.

252. Chung SJ, Kim D, Song JH, et al. Comparison of detection and miss rates of narrow band imaging, flexible spectral imaging chromoendoscopy and white light at screening colonoscopy: a randomised controlled back-to-back study. *Gut.* 2014;63(5):785-791.
253. Winawer SJ, Zauber AG, O'Brien MJ, et al. Randomized comparison of surveillance intervals after colonoscopic removal of newly diagnosed adenomatous polyps. The National Polyp Study Workgroup. *N Engl J Med.* 1993;328(13):901-906.
254. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med.* 1993;329(27):1977-1981.
255. Gargiullo PM, Rothenberg RB, Wilson HG. Confidence intervals, hypothesis tests, and sample sizes for the prevented fraction in cross-sectional studies. *Stat Med.* 1995;14(1):51-72.
256. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. *Am J Epidemiol.* 1974;99(5):325-332.
257. Kaminski MF, Bretthauer M, Zauber AG, et al. The NordICC Study: rationale and design of a randomized trial on colonoscopy screening for colorectal cancer. *Endoscopy.* 2012;44(7):695-702.
258. Giorgi Rossi P, Vicentini M, Sacchettini C, et al. Impact of Screening Program on Incidence of Colorectal Cancer: A Cohort Study in Italy. *Am J Gastroenterol.* 2015;110(9):1359-1366.
259. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of Benefits, Burden, and Harms of Colorectal Cancer Screening Strategies: Modeling Study for the US Preventive Services Task Force. *JAMA.* 2016.
260. Winawer SJ, Fischer SE, Levin B. Evidence-Based, Reality-Driven Colorectal Cancer Screening Guidelines: The Critical Relationship of Adherence to Effectiveness. *JAMA.* 2016;315(19):2065-2066.
261. Levin TR, Jensen CD, Zhao W, et al. 828 Adherence to Annual Fecal Immunochemical Testing in a Large, Community-Based Population. *Gastroenterology.* 2013;144(5):S-145-S-146.
262. Burch JA, Soares-Weiser K, St John DJ, et al. Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: a systematic review. *J Med Screen.* 2007;14(3):132-137.
263. Meester RG, Doubeni CA, Lansdorp-Vogelaar I, et al. Variation in Adenoma Detection Rate and the Lifetime Benefits and Cost of Colorectal Cancer Screening: A Microsimulation Model. *Jama.* 2015;313(23):2349-2358.
264. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Results of screening colonoscopy among persons 40 to 49 years of age. *N Engl J Med.* 2002;346(23):1781-1785.
265. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med.* 1993;328(19):1365-1371.
266. 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):1238-1252.
267. Doubeni C, Jensen C, Fedewa S, et al. Mailed fecal immunochemical test positivity and sensitivity vary with ambient temperature: a population-based study. *Journal of the American Board of Family Medicine* 2016(In press).
268. 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):1680-1688.

269. Gohel TD, Burke CA, Lankaala P, et al. Polypectomy rate: a surrogate for adenoma detection rate varies by colon segment, gender, and endoscopist. *Clin Gastroenterol Hepatol.* 2014;12(7):1137-1142.
270. Williams JE, Holub JL, Faigel DO. Polypectomy rate is a valid quality measure for colonoscopy: results from a national endoscopy database. *Gastrointest Endosc.* 2012;75(3):576-582.
271. Warren JL, Klabunde CN, Mariotto AB, et al. Adverse events after outpatient colonoscopy in the Medicare population. *Ann Intern Med.* 2009;150(12):849-857, W152.
272. Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR, Neugut AI. Risk of perforation after colonoscopy and sigmoidoscopy: a population-based study. *J Natl Cancer Inst.* 2003;95(3):230-236.
273. Robertson DJ, Lieberman DA, Winawer SJ, et al. Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. *Gut.* 2014;63(6):949-956.
274. Farrar WD, Sawhney MS, Nelson DB, Lederle FA, Bond JH. Colorectal cancers found after a complete colonoscopy. *Clin Gastroenterol Hepatol.* 2006;4(10):1259-1264.
275. Adler A, Wegscheider K, Lieberman D, et al. Factors determining the quality of screening colonoscopy: a prospective study on adenoma detection rates, from 12,134 examinations (Berlin colonoscopy project 3, BECOP-3). *Gut.* 2013;62(2):236-241.
276. Jensen CD, Doubeni CA, Quinn VP, et al. Adjusting for Patient Demographics Has Minimal Effects on Rates of Adenoma Detection in a Large, Community-based Setting. *Clin Gastroenterol Hepatol.* 2015;13(4):739-746.
277. 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):1752-1757.
278. 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):630-641.
279. U.S. Department of Labor. Bureau of Labor Statistics. 2014; http://www.bls.gov/. Accessed January 22, 2014.
280. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996:247-255.
281. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making.* 1985;5(2):157-177.
282. Rex DK, Schoenfeld PS, Cohen J, et al. Quality Indicators for Colonoscopy. *Am J Gastroenterol.* 2014.
283. Cooper GS, Xu F, Barnholtz Sloan JS, Schluchter MD, Koroukian SM. Prevalence and predictors of interval colorectal cancers in medicare beneficiaries. *Cancer.* 2012;118(12):3044-3052.
284. Hewett DG, Rex DK. Improving colonoscopy quality through health-care payment reform. *Am J Gastroenterol.* 2010;105(9):1925-1933.
285. Ladabaum U, Levin Z, Mannalithara A, Brill JV, Bundorf MK. Colorectal testing utilization and payments in a large cohort of commercially insured US adults. *Am J Gastroenterol.* 2014;109(10):1513-1525.
286. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *N Engl J Med.* 2000;343(3):169-174.
287. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med.* 2000;343(3):162-168.

288. Lee TJ, Rutter MD, Blanks RG, et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. *Gut.* 2012;61(7):1050-1057.
289. Zorzi M, Senore C, Da Re F, et al. Quality of colonoscopy in an organised colorectal cancer screening programme with immunochemical faecal occult blood test: the EQuIPE study (Evaluating Quality Indicators of the Performance of Endoscopy). *Gut.* 2015;64(9):1389-1396.
290. Pohl H, Srivastava A, Bensen SP, et al. Incomplete Polyp Resection During Colonoscopy-Results of the Complete Adenoma Resection (CARE) Study. *Gastroenterology.* 2012.
291. Symonds EL, Osborne JM, Cole SR, Bampton PA, Fraser RJ, Young GP. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. *J Med Screen.* 2015.
292. Meester RG, Zauber AG, Doubeni CA, et al. Consequences of Increasing Time to Colonoscopy Examination Following Positive Result From Fecal Colorectal Cancer Screening Test. *Clin Gastroenterol Hepatol.* 2016.
293. 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. *Bull World Health Organ.* 2008;86(4):317-319.
294. National Institute for Public Health and Environment. Bowel cancer screening programme,. 2014; http://www.rivm.nl/en/Topics/B/Bowel_cancer_screening_programme. Accessed Sep 28, 2016.
295. Kaminski MF, Anderson J, Valori R, et al. Leadership training to improve adenoma detection rate in screening colonoscopy: a randomised trial. *Gut.* 2016;65(4):616-624.
296. Ward ST, Hancox A, Mohammed MA, et al. The learning curve to achieve satisfactory completion rates in upper GI endoscopy: an analysis of a national training database. *Gut.* 2016.
297. Partin MR, Burgess DJ, Burgess JF, Jr., et al. Organizational predictors of colonoscopy follow-up for positive fecal occult blood test results: an observational study. *Cancer Epidemiol Biomarkers Prev.* 2015;24(2):422-434.
298. Chubak J, Garcia MP, Burnett-Hartman AN, et al. Time to Colonoscopy after Positive Fecal Blood Test in Four U.S. Health Care Systems. *Cancer Epidemiol Biomarkers Prev.* 2016;25(2):344-350.
299. Peterson K, Carson S, Humphrey L, Helfand M. Patients with Positive Screening Fecal Occult Blood Tests: Evidence Brief on the Relationship Between Time Delay to Colonoscopy and Colorectal Cancer Outcomes. *VA-ESP Project #09-199 (http://www.hsrd.research.va.gov/publications/esp/fecaloccult.cfm)* 2013.
300. Gellad ZF, Almirall D, Provenzale D, Fisher DA. Time from positive screening fecal occult blood test to colonoscopy and risk of neoplasia. *Dig Dis Sci.* 2009;54(11):2497-2502.
301. Parente F, Marino B, DeVecchi N, et al. Faecal occult blood test-based screening programme with high compliance for colonoscopy has a strong clinical impact on colorectal cancer. *Br J Surg.* 2009;96(5):533-540.
302. Parente F, Vailati C, Boemo C, et al. Improved 5-year survival of patients with immunochemical faecal blood test-screen-detected colorectal cancer versus non-screening cancers in northern Italy. *Dig Liver Dis.* 2015;47(1):68-72.
303. Brenner H, Altenhofen L, Katalinic A, Lansdorp-Vogelaar I, Hoffmeister M. Sojourn time of preclinical colorectal cancer by sex and age: estimates from the German national screening colonoscopy database. *Am J Epidemiol.* 2011;174(10):1140-1146.
304. Wattacheril J, Kramer JR, Richardson P, et al. Lagtimes in diagnosis and treatment of colorectal cancer: determinants and association with cancer stage and survival. *Aliment Pharmacol Ther.* 2008;28(9):1166-1174.

305. Ramos M, Esteva M, Cabeza E, Llobera J, Ruiz A. Lack of association between diagnostic and therapeutic delay and stage of colorectal cancer. *Eur J Cancer.* 2008;44(4):510-521.
306. Korsgaard M, Pedersen L, Sorensen HT, Laurberg S. Delay of treatment is associated with advanced stage of rectal cancer but not of colon cancer. *Cancer Detect Prev.* 2006;30(4):341-346.
307. Murchie P, Raja EA, Brewster DH, et al. Time from first presentation in primary care to treatment of symptomatic colorectal cancer: effect on disease stage and survival. *Br J Cancer.* 2014;111(3):461-469.
308. Wu GH, Wang YM, Yen AM, et al. Cost-effectiveness analysis of colorectal cancer screening with stool DNA testing in intermediate-incidence countries. *BMC Cancer.* 2006;6:136.
309. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, et al. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. *Ann Intern Med.* 2014;161(2):104-112.
310. Efron B, Tibshirani R. *An introduction to the bootstrap.* Vol 2nd. Boca Raton, Florida: CRC Press LLC; 1993.
311. Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. *Ann Intern Med.* 2014;160(3):171.
312. Rozen P, Liphshitz I, Barchana M. The changing epidemiology of colorectal cancer and its relevance for adapting screening guidelines and methods. *Eur J Cancer Prev.* 2011;20(1):46-53.
313. Shellnut JK, Wasvary HJ, Grodsky MB, Boura JA, Priest SG. Evaluating the age distribution of patients with colorectal cancer: are the United States Preventative Services Task Force guidelines for colorectal cancer screening appropriate? *Dis Colon Rectum.* 2010;53(1):5-8.
314. Ko CW, Sonnenberg A. Comparing risks and benefits of colorectal cancer screening in elderly patients. *Gastroenterology.* 2005;129(4):1163-1170.
315. Lin OS, Kozarek RA, Schembre DB, et al. Screening colonoscopy in very elderly patients: prevalence of neoplasia and estimated impact on life expectancy. *JAMA.* 2006;295(20):2357-2365.
316. Ness RM, Holmes AM, Klein R, Dittus R. Utility valuations for outcome states of colorectal cancer. *Am J Gastroenterol.* 1999;94(6):1650-1657.
317. Bureau of Labor Statistics. Current Employment Statistics - CES (National). www.bls.gov/web/empsit/ceseesummary.htm. Accessed Jan 16, 2013.
318. Bureau of Labor Statistics. Consumer Price Index. 2013; www.bls.gov/cpi/. Accessed Sep 23, 2013.
319. National Cancer Institute. SEER*Stat Software, version 5.3.1. Surveillance Research Program. . http://seer.cancer.gov/seerstat/. Accessed January 1, 2015.
320. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, et al. Individualizing colonoscopy screening by sex and race. *Gastrointest Endosc.* 2009;70(1):96-108, 108 e101-124.
321. Freedman AN, Slattery ML, Ballard-Barbash R, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. *J Clin Oncol.* 2009;27(5):686-693.
322. Painter J, Saunders DB, Bell GD, Williams CB, Pitt R, Bladen J. Depth of insertion at flexible sigmoidoscopy: implications for colorectal cancer screening and instrument design. *Endoscopy.* 1999;31(3):227-231.
323. Ladabaum U, Schoen RE. Post-Polypectomy Surveillance That Would Please Goldilocks-Not Too Much, Not Too Little, but Just Right. *Gastroenterology.* 2016;150(4):791-796.
324. Ransohoff DF, Lang CA, Kuo HS. Colonoscopic surveillance after polypectomy: considerations of cost effectiveness. *Ann Intern Med.* 1991;114(3):177-182.

325. Saini SD, Schoenfeld P, Vijan S. Surveillance colonoscopy is cost-effective for patients with adenomas who are at high risk of colorectal cancer. *Gastroenterology.* 2010;138(7):2292-2299, 2299 e2291.
326. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med.* 2012;366(8):687-696.
327. Chung SJ, Kim YS, Yang SY, et al. Five-year risk for advanced colorectal neoplasia after initial colonoscopy according to the baseline risk stratification: a prospective study in 2452 asymptomatic Koreans. *Gut.* 2011;60(11):1537-1543.
328. Laiyemo AO, Murphy G, Albert PS, et al. Postpolypectomy colonoscopy surveillance guidelines: predictive accuracy for advanced adenoma at 4 years. *Ann Intern Med.* 2008;148(6):419-426.
329. Lieberman DA, Weiss DG, Harford WV, et al. Five-year colon surveillance after screening colonoscopy. *Gastroenterology.* 2007;133(4):1077-1085.
330. Program SEaER. SEER Stat Fact Sheets: Colon and Rectum Cancer. 2015. Accessed 01/06, 2016.
331. National Center for Health Statistics. About the National Health Interview Survey. http://www.cdc.gov/nchs/nhis/about_nhis.htm. Accessed Oct 20, 2016.
332. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. *Ann Intern Med.* 2011;154(1):22-30.
333. Baxter NN, Warren JL, Barrett MJ, Stukel TA, Doria-Rose VP. Association between colonoscopy and colorectal cancer mortality in a US cohort according to site of cancer and colonoscopist specialty. *J Clin Oncol.* 2012;30(21):2664-2669.
334. Singh H, Nugent Z, Demers AA, Kliewer EV, Mahmud SM, Bernstein CN. The reduction in colorectal cancer mortality after colonoscopy varies by site of the cancer. *Gastroenterology.* 2010;139(4):1128-1137.
335. Brenner H, Hoffmeister M, Arndt V, Stegmaier C, Altenhofen L, Haug U. Protection from right- and left-sided colorectal neoplasms after colonoscopy: population-based study. *J Natl Cancer Inst.* 2010;102(2):89-95.
336. Chan AO, Jim MH, Lam KF, et al. Prevalence of colorectal neoplasm among patients with newly diagnosed coronary artery disease. *JAMA.* 2007;298(12):1412-1419.
337. Royce TJ, Hendrix LH, Stokes WA, Allen IM, Chen RC. Cancer screening rates in individuals with different life expectancies. *JAMA Intern Med.* 2014;174(10):1558-1565.
338. Joseph DA, Meester RG, Zauber AG, et al. Colorectal cancer screening: Estimated future colonoscopy need and current volume and capacity. *Cancer.* 2016.
339. Lieberman D. Screening for colorectal cancer in individuals at average risk: current methods and emerging issues. *JAMA Intern Med.* 2014;174(1):10-11.
340. American Gastroenterological Association. GI Quality Measures for 2017 Released in MACRA Final Rule. 2016; http://www.gastro.org/news_items/gi-quality-measures-for-2017-are-released-in-macra-final-rule. Accessed November 11, 2016.
341. Ladabaum U, Levin Z, Brill JV, Mannalithara A, Bundorf MK. How much are we paying for colonoscopy? Utilization rates and payments for colorectal testing 21 million commercial insurance enrollees in the United States. *Gastroenterology.* 2013;144(5):S1-S1136.
342. Fisher DA, Jeffreys A, Coffman CJ, Fasanella K. Barriers to full colon evaluation for a positive fecal occult blood test. *Cancer Epidemiol Biomarkers Prev.* 2006;15(6):1232-1235.

343. Powell AA, Nugent S, Ordin DL, Noorbaloochi S, Partin MR. Evaluation of a VHA collaborative to improve follow-up after a positive colorectal cancer screening test. *Med Care*. 2011;49(10):897-903.
344. Devi S. Obama vows to address veteran health-care scandal. *Lancet*. 2014;384(9946):841.
345. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut*. 2015;64(8):1327-1337.
346. Cho H, Klabunde CN, Yabroff KR, et al. Comorbidity-adjusted life expectancy: a new tool to inform recommendations for optimal screening strategies. *Ann Intern Med*. 2013;159(10):667-676.
347. Kahi CJ, Myers LJ, Slaven JE, et al. Lower endoscopy reduces colorectal cancer incidence in older individuals. *Gastroenterology*. 2014;146(3):718-725 e713.
348. van Hees F, Saini SD, Lansdorp-Vogelaar I, et al. Personalizing Colonoscopy Screening for Elderly Individuals Based on Screening History, Cancer Risk, and Comorbidity Status Could Increase Cost Effectiveness. *Gastroenterology*. 2015.
349. Shaukat A, Parekh M, Lipscomb J, Ladabaum U. Can calcium chemoprevention of adenoma recurrence substitute or serve as an adjunct for colonoscopic surveillance? *Int J Technol Assess Health Care*. 2009;25(2):222-231.
350. O'Connor A. So Will Processed Meat Give You Cancer? *New York Times*. Oct 31, 2015, 2015;SundayReview.
351. van Dam L, Kuipers EJ, Steyerberg EW, van Leerdam ME, de Beaufort ID. The price of autonomy: should we offer individuals a choice of colorectal cancer screening strategies? *Lancet Oncol*. 2013;14(1):e38-46.
352. Lieberman D. USPSTF CRC Guidelines Are Incomplete 2016; http://www.gastro.org/news_items/uspstf-crc-guidelines-are-incomplete. Accessed July 22, 2016.
353. Belderbos TD, Grobbee EJ, van Oijen MG, et al. Comparison of cecal intubation and adenoma detection between hospitals can provide incentives to improve quality of colonoscopy. *Endoscopy*. 2015;47(8):703-709.
354. Tosteson AN, Beaber EF, Tiro J, et al. Variation in Screening Abnormality Rates and Follow-Up of Breast, Cervical and Colorectal Cancer Screening within the PROSPR Consortium. *J Gen Intern Med*. 2016;31(4):372-379.
355. Lairson DR, Dicarlo M, Deshmuk AA, et al. Cost-effectiveness of a standard intervention versus a navigated intervention on colorectal cancer screening use in primary care. *Cancer*. 2014;120(7):1042-1049.
356. Myers RE, Sifri R, Daskalakis C, et al. Increasing colon cancer screening in primary care among african americans. *J Natl Cancer Inst*. 2014;106(12).
357. Kempe KL, Shetterly SM, France EK, Levin TR. Automated phone and mail population outreach to promote colorectal cancer screening. *Am J Manag Care*. 2012;18(7):370-378.
358. Corley DA, Jensen CD, Marks AR. Can we improve adenoma detection rates? A systematic review of intervention studies. *Gastrointest Endosc*. 2011;74(3):656-665.
359. Stock C, Uhlmann L, Hoffmeister M, Laux G, Kieser M, Brenner H. Identification of physicians with unusual performance in screening colonoscopy databases: a Bayesian approach. *Gastrointest Endosc*. 2015(1097-6779 (Electronic)).
360. de Wijkerslooth TR, de Haan MC, Stoop EM, et al. Burden of colonoscopy compared to non-cathartic CT-colonography in a colorectal cancer screening programme: randomised controlled trial. *Gut*. 2012;61(11):1552-1559.

361. Wilschut JA, Habbema JD, van Leerdam ME, et al. Fecal occult blood testing when colonoscopy capacity is limited. *J Natl Cancer Inst.* 2011;103(23):1741-1751.
362. U.S. Preventive Services Task Force. Screening for Colorectal Cancer: Recommendation Statement. 2008; http://uspreventiveservicestaskforce.org/uspstf08/colocancer/colors.htm. Accessed 06/05, 2013.
363. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int J Cancer.* 2002;99(2):260-266.
364. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med.* 2000;343(2):78-85.
365. Jiao S, Peters U, Berndt S, et al. Estimating the heritability of colorectal cancer. *Hum Mol Genet.* 2014;23(14):3898-3905.
366. Jensen CD, Quinn VP, Doubeni CA, et al. Colonoscopy Delay After a Positive Fecal Test and Risk of Colorectal Cancer-Related Outcomes. *Gastroenterology.* 2016;150(4):S45.
367. Goede SL, Rabeneck L, Lansdorp-Vogelaar I, et al. The impact of stratifying by family history in colorectal cancer screening programs. *Int J Cancer.* 2015;137(5):1119-1127.
368. Sachdeva M, Sachdeva N, Pal M, et al. CRISPR/Cas9: molecular tool for gene therapy to target genome and epigenome in the treatment of lung cancer. *Cancer Gene Ther.* 2015;22(11):509-517.
369. Fesnak AD, June CH, Levine BL. Engineered T cells: the promise and challenges of cancer immunotherapy. *Nat Rev Cancer.* 2016;16(9):566-581.

Appendices

Appendix A

Brief summary (English)

OPTIMIZING OUTCOMES OF COLORECTAL CANCER SCREENING

Colorectal cancer is a leading cause of death in the Netherlands and many other countries. Globally, on average one in 77 people die from the disease; in Europe, this proportion is even one in 37. Screening is an established means for reducing colorectal cancer mortality, and is implemented in an increasing number of countries around the world.

In this thesis, a number of studies were presented with the overarching aim to advance the knowledge on what defines optimal screening programs. Screening optimization is understood here as the maximization of program effects on population health, the minimization of adverse side-effects, and the containment of screening costs. Parameters considered for optimization included the test used for screening (endoscopic or stool-based tests), screening adherence rates, the time from a positive stool-based test result to follow-up examination with colonoscopy, the quality of colonoscopy, appropriate stopping ages for screening, and the intensity of screening. Besides a general introduction and discussion part, this thesis consists of nine studies presented over three parts. All presented studies were the product of international collaborations, and used advanced modeling techniques for the evaluation and optimization of screening.

Part I contains, in addition to a background chapter on screening and public health, a description of the Microsimulation Screening Analysis (MISCAN) model as developed by the Department of Public Health, Erasmus MC, Rotterdam, Netherlands. As was described, the model uses publicly available data on the prevalence of benign precursor polyps (adenomas) and the incidence and survival of colorectal cancer to simulate the natural history of the disease. The effects of screening are simulated based on the best available experimental data. The model has been validated multiple times, and is being used, amongst others, for the planning and monitoring of the national colorectal cancer screening program in the Netherlands.

Part II contains the results from independent collaborations with the U.S. Centers for Disease Control and Prevention and the American Cancer Society. The central question for this part was two-fold: first, what benefits may be expected from an increased screening participation in the United States, and second, whether sufficient capacity exists for higher screening uptake. In Chapter 3, MISCAN was used to estimate the potential public health impact of achieving an 80% uptake of screening in the target population by 2018. It was estimated that achieving this aim may prevent 277,000 new colorectal cancer cases and 203,000 disease-related deaths through 2030. In 2030 alone, this would be a 22% reduction in incidence and a 33% reduction in mortality. The findings from this study were used to inform the National Colorectal Cancer Roundtable campaign to increase screening uptake in the United States to

80% by 2018 (for example, see The *New York Times* of Oct 31, 2015) In Chapter 4, the predicted required colonoscopy capacity for achieving this aim was compared to the estimated available capacity in 2012. This study combined model predictions and survey data to find that there would be sufficient colonoscopy capacity for a hypothetical organized nation-wide screening program with primary colonoscopy or stool-based testing. With continuation of current test utilization patterns as estimated from National Health Interview Survey data, there would be less margin in 2012 capacity, but the expected colonoscopy demand would still remain within capacity limits. In Chapter 5, the proportion of colorectal cancer deaths attributable to nonuse of colorectal cancer screening was estimated for the United States, as part of a special issue of *Annals of Epidemiology* devoted to population attributable risk. MISCAN suggested that approximately 60% of deaths could have been prevented with full uptake of screening by the screening eligible population.

In Part III, the central aim was to estimate the impact of variation in several performance indicators of screening programs on program outcomes. In Chapter 6, previously unpublished data was used to compare the effectiveness of two important test modalities: colonoscopy screening versus annual sensitive fecal occult blood testing (sFOBT). The *National Colonoscopy Study* randomized approximately 3500 persons after their informed consent to compare adherence for colonoscopy and sFOBT over time. With the observed adherence, which was higher for colonoscopy (86%) than sFOBT (80% completing at least one test), the MISCAN model estimated that colonoscopy resulted in a substantially greater mortality reduction than sFOBT (59% versus 43%). In Chapter 7, another set of data was used to study the effects of variation in a physician-related performance indicator, the adenoma detection rate (ADR). The ADR is the proportion of a provider's screening colonoscopies detecting one or more adenomas. Kaiser Permanente Northern California collected detailed patient and physician-level information for almost 60,000 colonoscopies performed between 1998 and 2010, including post-colonoscopy cancer incidence. This information was used in MISCAN to estimate the variation in the sensitivity of colonoscopy for adenomas and cancer underlying the ADR variation. These estimates were used subsequently to estimate variation in long-term effects and cost of screening with colonoscopy. The results, which have been published in a 2015 issue of the *Journal of the American Association*, show that for every 5% point increase in ADR, colorectal cancer mortality decreased 13%, while adverse side-effects increased 10% due to more frequent referral for colonoscopy surveillance. The estimated net costs of screening and treatment were estimated to even decrease by 3% per 5% point ADR increase. In a follow-up study in Chapter 8, the influence of ADR on effects of screening with fecal immunochemical testing (FIT) was investigated. FIT screening requires colonoscopy for follow-up examination of positive test results and for colonoscopy

surveillance in adenoma patients. Although MISCAN suggested a substantial impact of ADR variation on health outcomes of FIT, the association was weaker for FIT than estimated for primary colonoscopy screening. In Chapter 9, another quality aspect was investigated in a more exploratory study on the consequences of increasing time from a positive FIT result to follow-up examination with colonoscopy. From the literature it is known that many people, particularly in settings with no organized screening programs, do not receive a diagnostic exam within 6 months from a positive fecal colorectal cancer test result. MISCAN estimated that longer times to follow-up of up to 12 months may increase the risk of cancer, and decrease overall screening benefits in terms of life-years gained by almost 10%.

The final research part, Part IV, focused on cost-effectiveness of screening in unscreened elderly people (75+ years) on the one hand, and cost-effectiveness of more frequent colonoscopy check-ups of patients with removed adenomas (surveillance) on the other hand. In Chapter 10, the cost-effectiveness of a single screening with colonoscopy, flexible sigmoidoscopy, or FIT was evaluated for unscreened elderly people with respect to their overall health status. MISCAN suggested that in people with poor overall health (severe comorbidities) screening may be warranted through age 80 years with a cost-effectiveness threshold of US$ 100,000 per quality-adjusted life-year gained. For healthy persons (no comorbidities), this was even up to age 86 years. In Chapter 11, finally, MISCAN was used to simulate competing strategies from a planned randomized clinical trial on colonoscopy surveillance in adenoma patients, the *European Polyp Surveillance* (EPoS) study. The EPoS study is designed to separately evaluate more versus less intensive surveillance in patients with low-risk adenomas (1-2 'tubular' adenomas up to 5mm in diameter) or high-risk adenomas (3-10 adenomas of any size and histology, or 1-10 adenomas ≥1cm). Modeling suggested that more intensive surveillance with 3-5 year intervals is cost-effective in the long-term compared to less intensive surveillance with 5-10 yearly colonoscopy or screening as recommended for average-risk individuals. Costs per life-year gained were generally less than US $25,000.

In the general discussion of Part V, several recommendations were made for future practice. First, it was recommended that screening is promoted as a very effective public health intervention under the condition that potential participants are adequately informed on both potential benefits and harms. It was further recommended that quality assurance programs are implemented in all screening settings, where the first important aim should be to measure a variety of relevant quality indicators. As we showed in this thesis, these included screening adherence rates, time factors (time to diagnostic follow-up), and indicators of colonoscopy quality. ADRs should be used in awareness of potential adverse effects of an exclusive focus on ADR as the primary colonoscopy quality metric. Finally, it was recommended

that screening is considered beyond age 75 years in patients who were not previously screened, and that 3-5 yearly colonoscopy surveillance is offered to patients with high- and low-risk adenomas, respectively. For effective surveillance policies it is deemed important that physicians assess adenoma size and multiplicity as accurately as possible.

More research was recommended to assess whether offering colonoscopy as the preferred method for screening followed by stool-based testing in case of refusal is an effective and cost-effective strategy to exploit the advantages of both screening methods. Other recommended areas for future research included the all-cause mortality effects of screening, the health impact of policies to stimulate higher ADRs, the search for novel DNA markers and development of improved non-invasive screening tests, and finally, the broad area of personalization in screening. Promising results from studies conducted within our department suggest that the effectiveness and efficiency gains from more personalized (or risk-stratified) screening may be substantial.

Colorectal cancer screening is a rapidly developing sub-discipline of public health. Innovations in screening and treatment will likely continue to change the screening landscape toward future years. This thesis ended on the optimistic note that with expected progress in the understanding of the etiology of colorectal cancer, the prognosis for future colorectal cancer patients will likely also improve.

Appendix B

Brief summary (Dutch)

HET OPTIMALISEREN VAN DE UITKOMSTEN VAN DARMKANKERSCREENING

Darmkanker is een belangrijke doodsoorzaak in Nederland en veel andere landen. Wereldwijd sterft gemiddeld ongeveer een op de 77 mensen aan de gevolgen van darmkanker; in Europa is dit zelfs een op de 37 mensen. Screening is een bewezen effectief middel ter voorkoming van darmkankersterfte en wordt daarom wereldwijd in steeds meer landen geïmplementeerd.

In dit proefschrift is een aantal studies gepresenteerd waarin is getracht een bijdrage te leveren aan het optimaliseren van screeningsprogramma's voor darmkanker. Optimalisatie is hierbij opgevat als het maximaliseren van de beoogde gezondheidseffecten, het minimaliseren van schadelijke neveneffecten, en het beperken van programmakosten. Onderzochte parameters voor optimalisatie zijn, de gebruikte screeningstest (inwendig of ontlastingsonderzoek), het deelnamepercentage aan screening, de tijd tussen een positieve ontlastingstest en inwendig vervolgonderzoek met 'coloscopie', de kwaliteit van het inwendig onderzoek, de stopleeftijd voor screening, en de intensiteit van screening. Naast een algemene inleiding en discussie bestaat dit proefschrift uit negen studies gepresenteerd over drie delen. Alle hier gepresenteerde studies zijn het resultaat van internationale samenwerkingen, en zijn uitgevoerd met behulp van een geavanceerd model voor de evaluatie en optimalisatie van darmkankerscreening.

In het eerste deel van dit proefschrift is naast een algemene inleiding op darmkankerscreening en volksgezondheid in het algemeen, het Microsimulatie-Screening-Analyse (MISCAN) model beschreven zoals ontwikkeld door de afdeling Maatschappelijke Gezondheidszorg, Erasmus MC, Rotterdam. Toegelicht is dat het model gegevens gebruikt over de prevalentie van goedaardige poliepen (adenomen) en de incidentie en overlevingskansen van darmkanker, om zo een inschatting te kunnen maken van het natuurlijk beloop van de ziekte. De effecten van screening hierop zijn gemodelleerd naar de bevindingen van de meest toonaangevende (gerandomiseerde) studies op dit gebied. Het model is meervoudig gevalideerd, en wordt onder andere gebruikt voor de planning en monitoring van het landelijke bevolkingsonderzoek naar darmkanker in Nederland.

Het tweede deel van dit proefschrift beschrijft de resultaten van samenwerking met respectievelijk de *U.S. Centers for Disease Control and Prevention* en de *American Cancer Society*. De centrale vraag voor dit gedeelte was tweeledig: allereerst is gekeken naar de potentiële baten voor de volksgezondheid van een hogere deelname aan screening in de Verenigde Staten, en ten tweede, naar de beschikbare capaciteit voor een dergelijke hogere deelname. In Hoofdstuk 3 is met MISCAN een schatting gedaan van de mogelijke maatschappelijke gezondheidseffecten van het behalen van een

deelname van 80% in de doelpopulatie voor screening in de Verenigde Staten per 2018. Naar schatting kunnen met het behalen van dit doel, tot 2030, afgerond 277,000 nieuwe gevallen van darmkanker worden voorkomen en 203,000 sterfgevallen aan darmkanker. In 2030 zou dit een geschatte reductie van 22% betekenen in het aantal nieuwe gevallen van darmkanker en een 33% reductie in het geschatte aantal sterfgevallen. Met de bevindingen van dit onderzoek is een nationale campagne in de Verenigde Staten geïnformeerd die 80% deelname per 2018 tot speerpunt heeft (zie bijvoorbeeld de *New York Times* van 31 Oktober 2015). In hoofdstuk 4 is in vervolg hierop de verwachte benodigde coloscopiecapaciteit voor het behalen van dezelfde doelstelling afgezet tegen de geschatte aanwezige capaciteit anno 2012. Dit combinatieonderzoek van modelschattingen en enquêtegegevens suggereert dat er ruimschoots voldoende capaciteit is voor een hypothetisch, georganiseerd, nationaal screeningsprogramma in de Verenigde Staten met coloscopie of ontlastingstesten. Met huidige realistische gebruikspatronen van screening is er minder marge in de capaciteit anno 2012, maar blijft de geschatte vraag toch binnen bereik van de geschatte capaciteit. In Hoofdstuk 5 is ten slotte, in het kader van een speciale editie van het tijdschrift *Annals of Epidemiology* over populatie attributief risico, gekeken naar het aandeel van darmkankersterfte in de Verenigde Staten dat voorkomen had kunnen worden met een 100% deelname aan screening. MISCAN suggereert dat met 100% deelname afgerond 60% van de sterfgevallen aan darmkanker voorkomen had kunnen worden.

In het derde deel van dit proefschrift stond de vraag centraal naar de impact van variatie in een aantal belangrijke prestatie-indicatoren van screeningsprogrammas op de uitkomsten van screening. In hoofdstuk 6 zijn met behulp van niet eerder gepubliceerde data de effecten vergeleken van coloscopie screening versus jaarlijks sensitief fecaal occult bloed testen (sFOBT). De *National Colonoscopy Study* heeft ruim 3500 bereidwillige personen gerandomiseerd ter vergelijking van de deelname aan coloscopie en sFOBT over tijd. Met de waargenomen deelname, die hoger was voor coloscopie (86%) dan voor sFOBT (80% deed ten minste één sFOBT), resulteert coloscopiescreening volgens MISCAN in een substantieel grotere reductie in darmkankersterfte dan sFOBT (59% versus 43%). In hoofdstuk 7 is met behulp van nieuwe empirische data onderzoek gedaan naar de effecten van variatie in een artsgebonden prestatie-indicator, de adenoom detectieratio (ADR). De ADR is het percentage van screeningscolonospieën waarbij een arts een of meer adenomen vindt. *Kaiser Permanente Northern California* heeft voor dit onderzoek van bijna 60,000 coloscopieën verschillende persoonsgebonden en artsgebonden gegevens verzameld, inclusief gegevens met betrekking tot kankerdiagnoses na coloscopie. Op basis van deze gegevens is met MISCAN een schatting gemaakt van de variatie in sensitiviteit van coloscopie voor adenomen en kanker die ten grondslag ligt aan de

variatie in ADR. Deze variatie in sensitiviteit is vervolgens vertaald naar lange-termijn uitkomsten en kosten van screening met coloscopie. De resultaten, die in 2015 zijn gepubliceerd in the *Journal of the American Medical Association*, laten zien dat voor elke 5% punt toename in ADR, de effecten van screening op darmkankersterfte met circa 13% afnemen, terwijl de schadelijke neveneffecten door een hogere verwijzing voor surveillance van adenoompatiënten met circa 10% toenemen. De geschatte netto kosten van behandeling en screening namen zelfs af met 3%. In Hoofdstuk 8 is vervolgens ook gekeken naar de invloed van ADR op de effecten van screening met fecaal immunochemische tests (FIT). Bij FIT screening wordt coloscopie gebruikt voor het diagnostisch onderzoek bij een positief testresultaat en voor surveillance in patiënten met adenomen. Hoewel MISCAN ook hier een sterk verband met lange-termijn effecten liet zien, was dit verband zwakker dan bij primaire screening met coloscopie. In Hoofdstuk 9 is ten slotte een derde kwaliteitsaspect van screening belicht in een exploratief onderzoek naar het effect van tijd tussen een positief FIT resultaat en het inwendig vervolgonderzoek met coloscopie. Uit de literatuur is bekend dat veel mensen, zeker in settingen zonder georganiseerde screening, pas 6 of meer maanden na een positieve testuitslag een coloscopie krijgen. MISCAN schat dat bij een toenemend interval tussen positieve FIT en diagnostische coloscopie van tot 12 maanden, de effecten van screening in termen van gewonnen levensjaren tot bijna 10% kunnen afnemen.

In het vierde deel, stond de vraag centraal naar, enerzijds, de kosteneffectiviteit van screening voor niet eerder gescreende ouderen (75+), en anderzijds, de kosteneffectiviteit van extra darmonderzoeken in patiënten met verwijderde adenomen (surveillance). In Hoofdstuk 10 is rekening houden met de algehele gezondheidsstatus van de patiënt geschat tot welke leeftijd screening met respectievelijk coloscopie, sigmoidoscopie, of FIT kosteneffectief kan zijn. MISCAN suggereert dat zelfs in personen met een slechte gezondheidstoestand screening tot 80 jarige leeftijd kosteneffectief kan zijn bij een grens van US $100,000, terwijl dat in gezonde ouderen wel tot leeftijd 86 het geval kan zijn. In Hoofdstuk 11, ten slotte, zijn met behulp van MISCAN twee strategieën gesimuleerd uit een geplande gerandomiseerde studie, de *European Polyp Surveillance (EPoS) study*. De EPoS studie zal meer met minder intensieve surveillancestrategieën vergelijken voor patiënten met laag-risico adenomen (1-2 'tubulaire' adenomen tot 5mm in diameter) en patiënten met hoog-risico adenomen (3-10 adenomen van willekeurige grootte en histologie, of 1-10 adenomen ≥1cm). Ons onderzoek suggereert dat intensievere surveillance met 3-5 jarige intervallen zeer kosteneffectief is op de lange termijn vergeleken met minder intensieve surveillance, of vergeleken met screening zoals aanbevolen voor mensen zonder verhoogd risico op darmkanker.

Uit de algemene discussie en samenvatting in deel 5 is een aantal kernaanbevelingen naar voren gekomen. In de eerste plaats, is aanbevolen screening actief te (blijven) promoten als een zeer effectieve maatregel ter bevordering van de volksgezondheid, op voorwaarde dat patiënten goed geïnformeerd worden over zowel de voor- als nadelen. Verder adviseren wij dat kwaliteitsbewaking vast onderdeel moet zijn van alle screeningsprogramma's, met als eerste doel om de kwaliteit te meten zoals die tot uitdrukking komt in verscheidene indicatoren, zoals deelnamepercentages, tijdsfactoren, en indicatoren voor kwaliteit van coloscopie. ADRs moeten met terughoudendheid gebruikt worden als voornaamste kwaliteitsmaat voor coloscopie totdat meer duidelijk bestaat over het effect van programma's om de ADR te verhogen. Verder is geadviseerd dat screening overwogen wordt voor mensen ouder dan 75 jaar wanneer iemand niet eerder gescreend is, en dat 3-5 jaarlijkse surveillance met coloscopie wordt aangeboden aan patiënten met hoog- en laag-risico adenomen. Voor een effectief surveillanceprogramma is het van belang dat artsen de grootte en het aantal van adenomen zo goed mogelijk meten.

Toekomstig onderzoek moet uitwijzen of het aanbieden van coloscopie als voorkeursoptie voor screening en het reserveren van ontlastingstesten voor niet-deelnemers een effectieve en kosteneffectieve strategie is om de voordelen van beide tests te benutten. Verder is onderzoek nodig naar het effect van darmkankerscreening op de totale sterfte in de populatie, naar het effect van strategieën ter verbetering van ADRs, ter verkenning van nieuwe DNA markers en ontwikkeling van verbeterde ontlastingstesten, en ten slotte, op het gebied van personalisering (of risico-stratificering) in screening.

Darmkankerscreening is sterk in ontwikkeling. Innovaties op het gebied van screening en behandeling zullen de praktijk van darmkankerscreening naar verwachting sterk blijven veranderen in komende jaren. Met de verwachte vooruitgang in het begrip van de oorzaken van darmkanker, zal ook de prognose voor toekomstige patiënten naar verwachting steeds verder verbeteren.

Appendix C

List of all publications

1. **Meester RG**, Lansdorp-Vogelaar I, Ladabaum U. Cost-Effectiveness Of Recommended Surveillance For Patients With Colorectal Adenomas. *This thesis*.
2. **Meester RG**, Lansdorp-Vogelaar I, Church TR, Feld AD, Mills G, Corley DA, Doubeni CA, Allen JI, Jordan PA, O'Brien MJ, Winawer SJ, Zauber AG. Effectiveness of colonoscopy versus annual sensitive fecal occult blood screening for colorectal cancer; A microsimulation model based on National Colonoscopy Study data. *This thesis*.
3. **Meester RG**, Doubeni CA, Zauber AG, van Ballegooijen M, Corley DA, Lansdorp-Vogelaar I. Impact of Adenoma Detection Rates on the Benefits of Fecal Testing versus Colonoscopy for Colorectal Cancer. *This thesis*
4. Omidvari A, **Meester RG**, Lansdorp-Vogelaar I. Cost-effectiveness of Surveillance in GI practice. *Best Practice & Research: Clinical Gastroenterology*. 2016. In press.
5. Mehta SJ, Jensen CD, Quinn VP, Schottinger JE, Zauber AG, **Meester RG**, Laiyemo AO, Fedewa S, Goodman M, Fletcher RH, Levin TR, Corley DA, Doubeni CA. Race / Ethnicity and Adoption of a Population Health Management Approach to Colorectal Cancer Screening in a Community-Based Healthcare System. *J Gen Intern Med*. 2016 Jul 13. [Epub ahead of print]
6. **Meester RG**, Zauber AG, Doubeni CA, Jensen CD, Quinn VP, Helfand M, Dominitz JA, Levin TR, Corley DA, Lansdorp-Vogelaar I. Consequences of Increasing Time to Colonoscopy Examination After Positive Result From Fecal Colorectal Cancer Screening Test. *Clin Gastroenterol Hepatol*. 2016 Oct;14(10):1445-1451.e8.
7. Joseph DA, **Meester RG**, Zauber AG, Manninen DL, Winges L, Dong FB, Peaker B, van Ballegooijen M. Colorectal cancer screening: Estimated future colonoscopy need and current volume and capacity. *Cancer*. 2016 May 20. doi: 10.1002/cncr.30070. [Epub ahead of print]
8. van Hees F, Saini SD, Lansdorp-Vogelaar I, Vijan S, **Meester RG**, de Koning HJ, Zauber AG, van Ballegooijen M. Personalizing Colonoscopy Screening for Elderly Individuals Based on Screening History, Cancer Risk, and Comorbidity Status Could Increase Cost Effectiveness. *Gastroenterology*. 2015 Aug 4. pii: S0016-5085(15)01068-9, doi: 10.1053/j.gastro.2015.07.042.
9. **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 Jun 16;313(23):2349-58.
10. Kroep S, Lansdorp-Vogelaar I, Rubenstein JH, de Koning HJ, **Meester RG**, Inadomi JM, van Ballegooijen M. An Accurate Cancer Incidence in Barrett's Esophagus: A Best Estimate Using Published Data and Modeling. *Gastroenterology*. 2015 Sep;149(3):577-585.e4.

11. **Meester RG**, Doubeni CA, Zauber AG, Goede SL, Levin TR, Corley DA, Jemal A, Lansdorp-Vogelaar I. Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018. *Cancer*. 2015 Mar 12. doi: 10.1002/cncr.29336.
12. **Meester RG**, Doubeni CA, Lansdorp-Vogelaar I, Goede SL, Levin TR, Quinn VP, van Ballegooijen M, Corley DA, Zauber AG. Colorectal cancer deaths attributable to nonuse of screening in the United States. *Ann Epidemiol*. 2015 Mar;25(3):208-213. e1.
13. van Hees F, Habbema JD, **Meester RG**, Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG. Should colorectal cancer screening be considered in elderly persons without previous screening? A cost-effectiveness analysis. *Ann Intern Med*. 2014 Jun 3;160(11):750-9.
14. Joustra P, **Meester RG**, van Ophem H. Can Statisticians beat surgeons at the planning of operations? Empirical Economics. 2013 Jun;44(3);1697-718.

Appendix D

PhD portfolio

PhD training	Year	Workload (ECTS)
Courses		
Certificate for basic teaching. Desiderius School, Erasmus University, Rotterdam, Netherlands.	2015-2016	1.1
Individuele begeleiding	2016	0.2
Hoorcollege geven	2016	0.2
Teach the teacher	2015	0.7
Mentee program. Cancer Intervention and Surveillance Modeling Network, National Cancer Institute, Bethesda, MD, United States.	2016	0.4
M.S. in Health Sciences, specialization Genetic Epidemiology. Netherlands Institute for Health Sciences, Erasmus University, Rotterdam, Netherlands.	2013-2015	70
Genetic-Epidemiologic Research Methods (grade 7.0/10)	2015	4.3
Introduction to the analysis of next-generation sequencing (grade 9.0/10)	2015	1.4
Study Design (grade 8.6/10)	2015	4.3
Linux for Scientists (attended)	2015	0.6
Family Based Genetic Analyses (grade 8.5/10)	2014	0.7
Conceptual Foundation of Epidemiologic Study Design (grade 9.0/10)	2014	0.7
Biostatistical Methods II: Classical Regression Models (grade 8.2/10)	2014	4.3
Introduction to Public Health (attended)	2014	0.7
Causal Inference (attended)	2014	0.7
Causal Mediation Analysis (attended)	2014	0.7
History of Epidemiologic Ideas (attended)	2014	0.7
Health Economics (attended)	2014	0.7
SNPs and Human Diseases (attended)	2014	1.4
Advances in Genome-Wide Association Studies (grade 6.8/10)	2014	1.4
Development Research Proposal (attended)	2014	2.5
Research period PIN Health Sciences (grade *"very good"*)	2014	29.6
Oral Research Presentation (attended)	2014	1.4
Principles of Research in Medicine and Epidemiology (attended)	2013	0.7
Genome Wide Association Analysis (attended)	2013	1.4
Principles of Genetic Epidemiology (attended)	2013	0.7
Advances in Genomics Research (attended)	2013	0.4
Genomics in Molecular Medicine (attended)	2013	1.4
Biostatistical Methods I: Basic Principles (exempted)		5.7
English Language (exempted)		1.4
Introduction to Medical Writing (exempted)		1.1
Planning and Evaluation of Screening. Netherlands Institute for Health Sciences, Erasmus University, Rotterdam, Netherlands.	2013	1.4
Scientific Writing. Department of Public Health, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2013	1.4
Introduction to Dynamo-HIA. Department of Public Health, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2013	0.2

Presentations

Invited oral presentation: National Colorectal Cancer Roundtable meeting, Bethesda, MD. Public health impact of 80x2018 by state.	2016	1.0
Oral presentation (2x): Cancer Intervention and Surveillance Modeling Network meeting, Gaithersburg, MD.	2016	1.4
Oral presentation: Erasmus MC Department of Public Health, Rotterdam.	2016	1.0
Poster of distinction: Digestive Disease Week, San Diego, CA. (Published: *Gastroenterology,* Vol. 150, Issue 4, S751-S752).	2016	1.0
Invited oral presentation: National Colorectal Cancer Roundtable meeting, Bethesda, MD. Public health impact of achieving 80% colorectal cancer screening rates in the United States by 2018.	2015	1.0
Oral presentation (2x): Cancer Intervention and Surveillance Modeling Network meeting, Rockville, MD.	2015	1.4
Oral presentation: International Cancer Screening Network meeting, Rotterdam Netherlands.	2015	1.0
Poster presentation: International Cancer Screening Network meeting, Rotterdam Netherlands.	2015	1.0
Oral presentation (Junior investigator award): Digestive Disease Week, Washington DC. (Published; *Gastrointestinal Endoscopy, Vol. 81, Issue 5, AB181-AB182; Gastroenterology,* Vol. 148, Issue 4, S-190)	2015	1.0
Oral presentation: Digestive Disease Week, Chicago IL. (Published: *Gastroenterology,* Vol. 146, Issue 5, S-70-S-71).	2014	1.0
Oral presentation: Cancer Intervention and Surveillance Modeling Network meeting, Bethesda, MD. Modeling the Cost-Effectiveness of an Organized Fecal Immunochemical Test Screening Program.	2013	1.0
Oral presentation: Erasmus MC Department of Public Health, Rotterdam.	2013	1.0
Oral presentation: Digestive Disease Week, Orlando FL. (Published: *Gastroenterology,* Vol. 144, Issue 5, S-190).	2013	1.0
Oral presentation: Dutch Society for Gastroenterologists (NVGE) Meeting, Veldhoven, Netherlands.	2013	1.0

National and International conferences

Annual National Colorectal Cancer Roundtable meeting, Bethesda, MD.	2016	0.2
Cancer Intervention and Surveillance Modeling Network meeting, Gaithersburg, MD.	2016	0.2
Digestive Disease Week, San Diego, CA.	2016	1.0
Annual National Colorectal Cancer Roundtable meeting, Bethesda, MD.	2015	0.4
Cancer Intervention and Surveillance Modeling Network meeting, Rockville, MD.	2015	0.4
International Cancer Screening Network meeting	2015	0.7
Digestive Disease Week, Washington, DC.	2015	1.0
Annual meeting National Colorectal Cancer Screening Program, Utrecht, Netherlands.	2014	0.2
Digestive Disease Week, Chicago, IL.	2014	1.0
Cancer Intervention and Surveillance Modeling Network meeting, Bethesda, MD.	2013	0.4
Digestive Disease Week, Orlando, FL.	2013	1.0
Dutch Society for Gastroenterologists (NVGE) Meeting, Veldhoven, Netherlands.	2013	0.1

Teaching		
Teaching assistant: Choices in health care, Bachelor in Medicine, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2016	1.0
Teaching assistant: Mortality metrics, Bachelor in Medicine, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2015-2016	1.0
Project supervisor: Public Health field studies, Bachelor in Medicine, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2016	1.0
Reviewing bachelor essays, Bachelor in Medicine, Erasmus MC University Medical Center, Rotterdam, Netherlands.	2015	0.5
Other		
Stanford University exchange, Stanford, CA, United States. Collaboration with Uri Ladabaum MD, Professor in Gastroenterology (Chapter 11 of this thesis). Attendance of various seminars in Departments of Medicine, Epidemiology, Economic Policy Research, Health Policy and Research, and Management Science and Engineering. In-person meetings with professors from each of these departments.	2016	*1 month* (2.0)
Reviewer for several medical journals, including New England Journal of Medicine, Journal of the American Medical Association, Clinical Gastroenterology and Hepatology, British Medical Journal Open, Medicine, Journal of Clinical Gastroenterology, World Gastroenterology Journal, Eastern Mediterranean Health Journal (WHO)	2015-2016	2.0
	Total	103.0

Appendix E

About the author

Reinier Meester was born in 1983 in Hoorn, the Netherlands. He received his primary education at Oud Zandbergen in Huis ter Heide and secondary education at the Chr. Gymnasium in Utrecht. After completing high school in 2001, he traveled through Australia and South-East Asia for a year. Upon his return, he commenced studies in Econometrics and Operations Research at the University of Amsterdam. In 2007, he obtained a Bachelor of Science degree in Econometrics and Operations Research, and in 2009, he obtained a Master's degree in Econometrics (*cum laude*), with the final research project focusing on operation room planning within the Academic Medical Center in Amsterdam. Before joining Erasmus University in 2012, he worked successively as an analyst for Publistat Media Research in Amsterdam, Robeco Quantitative Strategies in Rotterdam, and Towers Watson Risk Consulting and Software in Amstelveen. He also studied Divinity part-time at the Free University in Amsterdam, where he completed a minor in 2010 (30 ECTS). He currently lives in Rotterdam with his wife, two children, and expected third.

Acknowledgments

This work rests on the shoulders of many researchers, of whom I can mention only some here. Immensely important are all those within Erasmus who contributed to the development of the microsimulation model (Chapter 2) used in the included studies. Iris, thank you for your constructivism, wit and patience, to mention only a few of your qualities. You have been a big support both mentally and professionally. Marjolein, thank you for your trust and our methodological discussions. You have made me feel at home within Erasmus from the start. Harry, I appreciate your humor, critical review of this thesis, and the many opportunities that you have granted me indirectly to develop myself and to present our work internationally. I sincerely enjoyed working with each of you over these years.

Also important are the people involved through PROSPR, in particular, Ann Zauber at Memorial Sloan Kettering Cancer Center, Douglas Corley at Kaiser Permanente Northern California, and Chyke Doubeni at the University of Pennsylvania. Ann, thank you for your generosity (not just pound cakes), eye for detail, and our walks in Pasadena where you learned me all about your family history and may have sparked an interest in history within myself (Chapter 1). Doug, thank you for your excellent management, sharp reviews and 'wordsmithing' of countless draft manuscript versions. I don't think you ever missed one. Chyke, thank you for your excellent manuscript framing and –in my experience– mentoring including several 3 a.m. sessions for you before the start of clinics. I am excited that we may be collaborating closer in the future on interesting new projects.

I would also like to thank Uri Ladabaum at Stanford University, for welcoming me to one of the finest research institutions. Thank you for your generosity in giving me the opportunity to meet a variety of faculty, our pleasant collaboration (Chapter 11), the travel tips for the weekends, and delicious lunch breaks (the Mexican food!). I cannot rule out ever returning to the Bay Area.

I would also like to extend my gratitude to the other model groups within CISNET, represented by Amy Knudsen at Massachusetts General Hospital, Carolyn Rutter at RAND, and Karen Kuntz at the University of Minnesota. I have come to appreciate your modeling expertise, our recent collaborations, and our, in my case, biennial meetings. I look forward to the next.

There are many other people that have played a role for this thesis, and should be acknowledged. First, within Erasmus, Frank van Hees (Chapter 2, 10), Luuk Goede, Miriam van der Meulen, Andrea Gini, Alex van der Steen, Steffie Naber, and Amir Omidvari. At Kaiser Permanente in California, Chris Jensen, Theodore (T.R.) Levin, Virginia Quinn, Joanne Schottinger, Nirupa Ghai, Wei Zhao, and Amy Marks (Chapter 5-9). At Memorial Sloan Kettering Cancer Center, Sidney Winawer, Sean Ryans, and Deborah Kuk, and, at the University of Minnesota, Tim Church (Chapter 6). At the American Cancer Society, Ahmedin Jemal, and Stacey Fedewa (Chapter 3). At the

Centers for Disease Control and Prevention, Djenaba Joseph and Laura Seeff, and, at Battelle, Fred Dong, Dianne Manninen, and Linda Winges (Chapter 4). At Veterans Affairs, Jason Dominitz and Mark Helfand (Chapter 9). It has been an enriching experience for me to work with so many different people.

I would like to thank and acknowledge everyone at NIHES, foremost Albert Hoffman, now head of Epidemiology at the Harvard School of Public Health, for their role in the excellent education program which I have been enabled to follow. I learned a lot and enjoyed particularly the guest lectures by speakers such as Kenneth Rothman, Alfredo Morabia and Miguel Hernan. You have been an inspiration. Thanks to Dr. Rothman for his help with some of the wording for the JAMA article (Chapter 7).

Thanks to the JAMA editorial staff for giving me the opportunity to revise our manuscript (Chapter 7), and raising the bar for me for what a modeling study should look like.

Thanks to the U.S. National Cancer Institute for supporting this work financially.

Finally, thanks for the indirect support from everyone dear to me, including but not limited to 'roomies' within Erasmus, friends, and family. Thanks mum and dad, brothers and sisters, family in law, for your love and support over the years. Anneke, Elianne, Stephan, "you are the sunshine of my life".

SDG