UNIVERSITY OF CALGARY

Economic Evaluation of Colorectal Cancer Screening for Average Risk Individuals

by

Steven James Heitman

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

MAY, 2012

UNIVERSITY OF
CALGARY

Abstract

Background: Colorectal cancer (CRC) is a common deadly cancer. Screening for CRC saves lives and is cost-effective, but it is unclear if one or more of the screening options are preferred.

Methods: A Markov model was developed and validated and then used to conduct an economic evaluation of CRC screening for average risk individuals. All current CRC screening modalities and up to date CRC treatment costs were considered. A systematic review and meta-analysis of CRC and adenomatous polyp prevalence was also performed to inform the model.

Results: The prevalence of non-advanced adenomas, advanced adenomas and CRC in 50-64 and 65-75 year olds was 17.1%, 3.8% and 0.1% and 17.3%, 8.2% and 0.7%, respectively. In the base case analysis CRC screening with annual FIT reduced the risk of CRC and CRC-related deaths and was associated with lower health care costs compared to no screening and the other screening options.

Conclusion: Health policy decision makers should prioritize funding for CRC screening using FIT.

Acknowledgements

Firstly, I would like to thank my wife and closest friend Nili for putting up with me while completing this second Master's Degree. Nili, life has certainly been eventful over the past couple of years! Thank you for being there for me. My two beautiful children Hilit and Joshua also deserve thanks for being so tolerant during the extra times when their daddy was away working towards this Master's.

I would like to thank Dr. Braden Manns for introducing me to health economics and for being such an incredible mentor for me over the past decade during both my medical training and this Master's. Braden, you have been tremendously supportive of my endeavours over the years and I simply cannot thank you enough. I would also like to thank Dr. Bob Hilsden for supervising me and for his efforts in connecting my research interests within the context of gastroenterology. Bob, I look forward to our continued collaboration. A thank you as well to my committee members, Dr. Elizabeth McGregor and Dr. Deborah Marshall for their important input on the project.

Paul Ronksley deserves recognition as well for his tireless efforts in completing the systematic review and meta-analysis (aka "the polyp extravaganza") with me and for making me laugh. My sincere gratitude also goes to Dr. Brenda Hemmelgarn for mentoring Paul and I through the systematic review.

Finally, I would like to extend a special thank you to Flora Au for her many hours of assistance with the Markov model. A model of this magnitude could not have been created without your tremendous help Flora – thank you!

The Canadian Association of Gastroenterology, Canadian Institutes of Health Research, and Alberta Innovates - Health Solutions (formerly Alberta Heritage Foundation for Medical Research) provided financial support.

Dedication

This thesis is dedicated in loving memory of my father, Dr. James Hubert Heitman.

Contributions of Authors

Prevalence of adenomas and colorectal cancer in average risk individuals: A systematic review and meta-analysis. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, Hemmelgarn BR. Clin Gastroenterol Hepatol. 2009 Dec;7(12):1272-8

SJ Heitman: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript

PE Ronksley: study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript

RJ Hilsden: analysis and interpretation of data, critical revision of the manuscript for important intellectual content

BJ Manns: analysis and interpretation of data, critical revision of the manuscript for important intellectual content

A Rostom: critical revision of the manuscript for important intellectual content, statistical analysis

BR Hemmelgarn: study concept and design, analysis and interpretation of data, critical revision of the manuscript for important intellectual content, statistical analysis, study supervision

Colorectal cancer screening for average-risk North Americans: An economic evaluation. Heitman SJ, Hilsden RJ, Au F, Dowden S, Manns BJ. PLoS Med 2010 Nov 23;7(11):e1000370.

SJ Heitman: designed and validated the model, performed the economic evaluation and wrote the manuscript

RJ Hilsden: helped develop the model, participated in drafting and revising the manuscript

F Au: assisted in the construction of the model and its validation, helped perform the economic evaluation and revised the manuscript

S Dowden: provided the costing data for the management of colorectal cancer and revised the manuscript

BJ Manns: supervised the creation of the model, provided overall supervision of the economic evaluation and revised the manuscript

Table of Contents

List of Symbols, Abbreviations and Nomenclature

If you do not have any symbols, abbreviations, or specific nomenclature in your thesis, you do not need to fill out this table. To add another row to the table, with your cursor in the bottom right cell, press the TAB key (beside the letter Q on your keyboard).

Symbol	Definition
ACRCSP	Alberta Colorectal Cancer Screening Program
CADTH	Canadian Agency for Drugs and Technologies in Heath
CBA	cost benefit analysis
CCSC	Colon Cancer Screening Centre
CEA	cost-effectiveness analysis
CRC	colorectal cancer
CTC	computed tomographic colonography
CUA	cost-utility analysis
FIT	fecal immunochemical test
FOBT	fecal occult blood test
HRQOL	health related quality of life
ICER	incremental cost-effectiveness ratio
QALY	quality adjusted life year
RCT	randomized controlled trial

CHAPTER ONE: BACKGROUND

Brief Introduction

Colorectal cancer (CRC) is a common deadly cancer. In addition, CRC is costly to society from the standpoint of lost quantity and quality of life and it is expensive to treat. However, CRC can be prevented through screening. Yet despite being recommended by numerous consensus guidelines in North America and Europe, the uptake of screening, particularly in Canada, is low. In today's fiscally strained environment, it is imperative that governments fund not only effective, but cost-effective interventions. CRC screening is felt to offer good value for health care money, but there are many available screening options and it remains unclear if one or more are preferred. It is unlikely that health care jurisdictions will be able to continue providing the necessary funding and infrastructure to support all of the currently available CRC screening modalities.

Economic evaluations can assist medical decision making. Although a number of economic evaluations assessing the impact and cost-effectiveness of CRC screening have been conducted, several important limitations exist with these previous studies. Given these limitations, we aimed to develop and validate a comprehensive decision analytic model that considered **all** relevant CRC screening strategies and recent CRC treatment costs so that the cost-effectiveness of CRC screening among average risk individuals could be rigorously assessed. During the initial construction of this model we discovered that precise prevalence rates for precursor adenomatous polyps and CRC among a truly average risk population had surprisingly not been reported. Instead, a vast number of

studies reporting a wide range of prevalence rates were available without a clear explanation for the observed differences. Yet, it seemed plausible that these baseline inputs could be critically important in driving the cost-effectiveness of the various CRC screening strategies and we had no justification for selecting a given study over the others. Systematic reviews facilitate the object appraisal and summary of evidence. Furthermore, meta-analysis is a power tool for pooling data to generate more precise parameter estimates and can be used to help account for heterogeneity in data. Systematic reviews and meta-analyses are often used to inform decision analytic models and one had not been done on the prevalence of adenomatous polyps and CRC among average risk individuals.

Thus, the primary objective of this thesis was to conduct an economic evaluation of CRC screening for average risk individuals using a decision analytic model that was developed and validated for this purpose. In order to inform the model a systematic review and meta-analysis of adenomatous polyp and CRC prevalence among average risk individuals was first performed. Both the systematic review and economic evaluation summarize and synthesize a vast body of literature and make up the research content of this thesis. Although each study has been published independently, these manuscripts together form a coherent body of research. Therefore, the manuscripts are presented together in the framework of a paper-based thesis.

Following the background and literature review contained in chapter one, the systematic review and meta-analysis [2] and economic evaluation [3] are presented in published manuscript form in chapters two and three, respectively. The fourth and final chapter provides a general discussion of the policy implications arising from this thesis

and a description of the future work that should follow. It is intended that the present work will help inform future CRC screening guidelines and drive health care decision making.

Background and Literature Review

1.0 The Burden of Colorectal Cancer

1.0.1 Epidemiology

In Canada colorectal cancer (CRC) is the fourth most commonly diagnosed cancer and the second leading cause of cancer-related death in both men and women [4]. In 2011, it was estimated that approximately 22,200 Canadians would be diagnosed with CRC and 8,900 would die from it [5].

The majority of CRCs are sporadic arising in individuals without increased familial risk. It is known that CRC is more common with increasing age and in those with a family history [6], in males [7], among smokers [8], and in those who are overweight [9]. However, prior to completing this thesis, precise estimates on the prevalence of precursor adenomatous polyps and CRC among the average risk population that comprise the majority of the disease burden were not known.

1.0.2 The Cost of Colorectal Cancer to the Health Care System and Society

In addition to its impact on loss of life, the economic burden of managing CRC is considerable. Canadian data are limited, but the lifetime direct health care costs of managing a Canadian cohort of nearly 17,000 patients with CRC was estimated at $520 million in 2000 [10]. A study evaluating the economic burden to Medicare alone in the US

estimated that that the costs of care in 2000 in the initial, continuing and last year of life phases of care were approximately $3.18 billion, $1.68 billion and $2.63 billion, respectively [11]. These authors also considered a current trends scenario with decreasing incidence, improving survival, and increasing costs and determined that these figures are expected to increase to $5.19 billion, $3.57 billion and $5.27 billion, respectively by 2020 [11].

The cost of managing CRC has increased due to advancements in chemotherapeutic agents and this will likely continue to drive much of the increase in costs in managing patients with CRC. The monoclonal antibody Bevacizumab has been shown to extend life by an average of 5 months in patients with metastatic disease [12]. A Canadian analysis from a direct-payer perspective estimated that the additional cost to achieve this benefit would be nearly $40,000 per patient in 2005 dollars [13].

The medical costs of treating patients with CRC typically have not taken into consideration the large losses in economic productively that result from premature death due to CRC. Using the human capital approach in which productivity is heavily influenced by working-age individuals and earnings, Bradley et al. [14] estimated that the present value of lifetime earnings lost from CRC in the US will be $12.8 billion in 2010. Another approach to estimating lost productivity that values each year of life lost equally ($150,000 per year) regardless of age, employment status, earnings, care-giving or housekeeping activity, estimated lost productivity due to CRC in the year 2000 in the US at just over $90 billion [11]. CRC is thus an important public health issue and is costly to society.

1.1 Screening for Colorectal Cancer

1.1.1 Adenoma-Carcinoma Sequence

It is widely accepted that most CRCs develop from pre-cancerous adenomatous tissue. Evidence for this is supported by epidemiologic [15], clinicopathologic [15, 16] and molecular genetic [17] studies. In the adenoma-carcinoma sequence [18] a series of genetic mutations accumulate over time resulting in normal colonic epithelium becoming dysplastic (adenomatous) and then malignant. During this process, morphological changes occur including the formation of the adenomatous polyp. These changes are generally believed to take place over a decade or more among those at average risk [19]. Polyps and cancers bleed intermittently and shed DNA into stool and they can generally be identified visually. These factors make CRC amenable to screening and form the basis for the available stool-based (guaiac-based fecal occult blood test (FOBT), fecal immunochemical test (FIT), and fecal DNA), radiological (barium enema and computed tomographic colonography (CTC)) and endoscopic screening strategies (sigmoidoscopy and colonoscopy). Information of the effectiveness of each of these screening modalities can be found in Appendix 1.

1.1.2 Clinical Practice Guidelines

Current clinical practice guidelines in both Canada [20] and the U.S. [21-23] recommend screening for CRC starting at age 50 in average risk individuals. In Canada average risk CRC screening is recommended until the age of 74 [20] whereas The U.S. Preventive Services Task Force recommends routine screening until the age of 75 [21].

Despite these recommendations, the uptake of CRC screening remains disappointingly low, especially in Canada at far less than 50% [24].

To date, no single screening modality has been shown to be uniformly superior and thus current clinical practice guidelines continue to recommend that patients be offered a choice between the available screening options. Differences exist, but in general the guidelines support FOBT or FIT every 1-2 years, sigmoidoscopy every 5-10 years and colonoscopy every 10 years. The American College of Gastroenterology [23] and the joint committee of the American Cancer Society, the U.S. Multi-Society Task Force on Colorectal Cancer and the American College of Radiology [22] support the use of CTC every 5 years. However, a number of issues including whether or not to report diminutive lesions (polyps < 5 mm not routinely reported by radiologists) and the impact of radiation exposure over time need clarification. The US Preventive Services Task Force does not endorse CTC citing that the "evidence is insufficient to assess the benefits and harms of the technology" [21]. Nevertheless, US Medicare and Medicaid do not cover CTC for CRC screening and it is not currently funded for primary screening in most jurisdictions in Canada.

The joint committee recently concluded that there are now sufficient data to include fecal DNA as an acceptable option for CRC screening, but the recommended interval remains uncertain [22]. However, the American College of Gastroenterology [23] and the U.S. Preventive Services Task Force [21] do not currently support fecal DNA for CRC screening and it has not been recommended in Canada. Nonetheless, the test is no longer commercially available.

1.2 Health Economics

1.2.1 Scarcity, Choice and Opportunity Cost

Resources whether people, time or the proportion of the provincial budget available for health care are scarce. Health care expenditures in general, and in particular for costly new drug treatments has risen exponentially over the past decade [25]. Given the expectation for further medical advancement together with an aging population, the cost of providing health care will undoubtedly continue its upward trajectory.

Given competing demands for scarce resources, choices must be made regarding which programs to fund. These decisions are rarely straightforward – e.g. should we create additional capacity in the intensive care unit or expand ambulatory clinics? Whatever choice is made, a decision to allocate scarce resources for a given program comes at a cost of not being able to allocate the same resources towards another potentially equally important use. Thus, the "opportunity cost" of any program is the value of the forgone benefits achievable by choosing to allocate resource to one program over another [1].

1.2.2 Economic Evaluation – An Overview

Economic evaluation facilitates decision making under conditions of scarcity when by definition one cannot produce all the desired outputs. Its analytical framework incorporates the concepts of scarcity, choice and opportunity cost. According to Drummond et al. [1], economic evaluation is the comparative analysis of alternative courses of action in terms of both their costs and consequences. There are four basic types of economic evaluation: cost-minimization analysis, cost-effectiveness analysis,

cost-utility analysis and cost-benefit analysis. Costs are measured similarly across all types of economic evaluation. However, the manner in which health outcomes are measured is unique to each type of analysis. A brief overview of each of the four types of economic evaluation is presented in the following sections. Additional detail can be found elsewhere [1].

1.2.3 Cost-Minimization Analysis

In a cost-minimization analysis the health outcomes of the treatments/programs under consideration are believed to be identical. Thus, the analysis becomes a simple comparison of the costs involved. For example, although a new migraine prevention medication was shown to be as effective as the standard therapy, it was determined to be the most costly treatment option in a cost-minimization analysis [26]. The authors concluded that the standard therapy offered better value for money unless additional benefits could be shown for the new medication. One issue that often arises in cost-minimization analysis is whether the treatments are in fact truly equivalent, as subtle differences in effectiveness, safety or convenience often exist.

1.2.4 Cost-Effectiveness Analysis

It is often the case that different interventions vary not only in cost, but also in their ability to achieve a given health outcome of interest. In a cost-effectiveness analysis (CEA), health outcomes are measured in a single common denominator, typically a naturally occurring unit such as the number of life-years saved. For example, accelerated tissue plasminogen activator (t-PA) was shown to have a cost per life-year saved of

approximately $33,000 compared to streptokinase, the traditional therapy at the time for patients suffering an acute myocardial infarction [27].

An important advantage of a CEA is that it can be used to compare any alternatives (even if very different) as long as they have a common effect. Open- heart surgery can be compared to mandatory seat belt legislation if the common health outcome is the number of life-years saved. However, a common unit of health benefit is not always available. For instance, it would be difficult to compare the health benefits of a renal transplantation program to that of a women's shelter. In addition, even when a common metric exists, the benefits of the health outcome in question are not always easy to comprehend. For instance, how could a cost of $10,000 per heart attack avoided be compared with a cost of $5,000 per additional central line insertion avoided to inform health care policy?

1.2.5 Cost-Utility Analysis

Health-related quality of life (HRQOL) is an important health outcome. Indeed, the greatest impact of many health programs rests in the ability to improve physical and social function and psychological well-being. Thus, CUA is appropriate when HRQOL is an important consideration. In a CUA a single composite index that combines length of life and quality of life termed the quality adjusted life year (QALY) is used as a common measure of health benefit [1]. This can facilitate comparisons across different interventions lacking a common unit of health benefit, which is an important advantage of this type of analysis.

A variety of tools have been developed to measure HRQOL using two general approaches: health status and health utility [28]. Health status is generally assessed using questionnaire-based instruments of which both generic and disease specific examples exist. In contrast, utility is a health state preference that ranges from zero (death) to one (perfect health), although values less than zero (worse than death) are also possible. Health utility values are what are used to produce QALYs. Utilities can be generated either directly or indirectly. The three direct methods include the standard gamble, the time trade-off and the visual analogue scale. The main disadvantages of these direct measurement techniques are that they are expensive and cumbersome to perform. A number of widely available indirect instruments are also available for obtaining utilities including the Health Utilities Index, the EQ-5D, and the SF-6D and SF-15D [28].

Different programs can have a wide range of health outcomes that are difficult to compare. By generating a common unit of output (the QALY), a CUA facilitates allocative decision making across different programs. For example, if it has been decided that a given level of funding will be spent to improve the health of patients with a certain medical condition, an intervention that costs $10,000 per QALY (intervention A) would likely be considered more attractive than another that costs $50,000 per QALY (intervention B). More health (QALYs) can be "purchased per dollar spent" with intervention A compared to intervention B.

1.2.6 Cost-Benefit Analysis

The last type of analysis is the cost-benefit analysis (CBA). In CBA the costs and consequences of alternative interventions are valued in monetary terms which has been

considered the broadest measure of value [1]. The advantage of this approach is that one can directly answer the question of whether the benefits of a therapy for example justify its costs without reference to another intervention or external benchmark. In contrast, the result of a CUA study can only be interpreted in relation to some standard that is considered 'acceptable' to pay for a QALY. Hence, CBA can directly address questions of allocative efficiency which the other methods cannot. CBA is used less frequently in health care due to the fact that many people find it difficult to value health effects such as quality of life in monetary terms [29]. However, one can estimate the monetary value of an intervention's effect on health by assessing an individual's willingness-to-pay for improved health.

1.2.7 Choosing a Type of Economic Evaluation for Colorectal Cancer Screening

Patients with CRC not only face obvious losses in quantity of life, but suffer significant reductions in their quality of life as well. As such, CUA is an appropriate method of economic evaluation of interventions for CRC including screening given that the QALY can capture these important changes in quantity and quality of life. Furthermore, interventions such as screening for many other chronic diseases and cancers that broadly impact health have been evaluated in this manner. Therefore, economic evaluations of CRC screening using CUA can facilitate comparisons across programs that are competing for scarce health care dollars.

1.3 Conducting an Economic Evaluation

Economic evaluations are conducted using two general approaches: 1) as part of a clinical trial or 2) through the use of decision analytical modeling. Both approaches are discussed in some detail in the following sections.

1.3.1 Alongside a Clinical Trial

The randomized controlled trial (RCT) is considered the most robust study design of clinical research. By controlling for both known and unknown confounders randomization helps ensure that the conclusions drawn from RCTs are valid. RCTs typically address topical issues and approval by licensing bodies often rests on the results of these pivotal studies. Economic data are becoming increasing important to funders and thus RCTs lend the opportunity to capture patient-specific data on costs and outcomes simultaneously. Therefore, performing an economic evaluation alongside a clinical trial can capitalize on efficiencies in data capture as well as timeliness.

There are also important disadvantages to conducting an economic evaluation alongside a clinical trial. Although bias is minimized, RCTs are not necessarily reflective of “real life” clinical practice. Feasibility is also a major factor, especially considering clinical trials of CRC screening which generally require tens of thousands of subjects with follow-up lasting more than a decade. Even when large clinical trials are performed, long-term outcomes and costs are frequently unavailable given that follow-up may halt after a desired clinical endpoint has been attained. Therefore, the long-term impact of an intervention on costs and health outcomes are not typically available from an RCT. Finally, surrogate (intermediate) endpoints are often selected over final outcomes to

minimize the duration and ultimately the cost of an RCT. How to incorporate health benefit which has only been measured using surrogate endpoints into an economic evaluation can be problematic. All of these issues make conducting an economic evaluation exclusively alongside RCTs impractical.

1.3.2 Decision Analysis

Decision making in the "real world" can rarely be achieved solely on the basis of a single data source such as an RCT. Human conditions are complex and long-term outcomes often depend on many uncertain variables. Decision analytical modeling has been developed as a framework to assist decision-making under conditions of uncertainty [1]. In this type of analysis a model is constructed as a simplification of reality in order to reproduce the health outcomes that are associated with a health condition. The model is then populated with data, typically from a number of sources. High quality clinical trial data are used if available, but other data sources are also generally required. Although often viewed by sceptics as artificial, non-transparent and susceptible to bias from the modeller [30], a major advantage of modeling is that sensitivity analyses can be performed where uncertain variables are varied to assess their influence on the results. Though it is true that a model can never be "real life", such a criticism is also relevant to the artificial conditions that accompany an RCT. On the other hand, a model can be used to overcome many of the shortcomings of an RCT, such as a relatively short time horizon. When designed and validated rigorously (see below) and when extensive sensitivity analyses are performed on variables where data are lacking, decision analytical models can produce high quality results.

1.3.3 Types of Decision Models

Two commonly used model designs are the simple decision tree and the Markov model [1]. In a simple tree structure, the decision in question is represented through a series of chance nodes that define points of transition from one event to another. Branches emanating from the chance nodes define the events, which occur at a given probability. The combination of the different branches in the tree establishes a series of outcomes that depend on the transition probabilities, which can be varied. Although widely used, simple decision trees are not particularly suitable for addressing more complex problems. For events that can recur over long time horizons, as is often the case of chronic conditions, the framework of a simple decision tree can become unmanageable.

In contrast, a Markov model has features designed to overcome the limitations of simple decision trees. Whereas different outcomes within a decision tree are created through alternative branches, events in a Markov model are based on transitions between a series of defined health ‘states’ that an individual can assume at any point in time. Time elapses at pre-set time intervals termed cycles, the length of which depends on the nature of the process being modeled. Movement can occur both within and between the states, but transitions can only occur once per cycle.

1.3.4 Analyzing Markov Models

The traditional method in which Markov models are analyzed is through cohort simulation. In cohort simulation a hypothetical cohort is distributed proportionally among the possible states with subsequent movement within and between states from cycle to cycle occurring according to the transition probabilities in the model. These proportions

are then linked to the costs and outcomes associated with each state in order to calculate the overall costs and outcomes for each of the alternatives being compared. In essence, this sorting of the cohort generates the average experience of the patients in the cohort [29].

Another approach to analyzing a Markov model is through Monte Carlo simulation. There are two types: first-order simulation (also called microsimulation) and second-order simulation. First-order simulation involves the use of random numbers to generate a single path (also called a trial) through the model, where transitions of higher probability events occur more frequently. Instead of proportions of patients making transitions through the model, individual patients are simulated going from cycle to cycle. First order Monte Carlo simulation can be used to track individual events within a model [29], which cannot be done with cohort simulation alone. First order Monte Carlo simulation can also be used to overcome a limiting property of a Markov model whereby a given cycle is not dependent on the history of previous cycles. Indeed, one of the challenges of analyzing Markov models involves accounting for important events that occurred prior to the current cycle and how this "history" (e.g. type, order of events) influences subsequent state transitions in the model. "Memory" can be incorporated into Markov models by simply creating additional states that reflect past experiences. Although effective, this can result in cumbersome models. Another method to overcome the inherent lack of memory in a Markov model is to perform first order Monte Carlo simulation using tracker variables that identify or "flag" each individual's characteristics and events experienced running through the model. Tracker variables can then be used to adjust subsequent transition probabilities, health state utilities and corresponding costs based on the individual's history through the model.

Second order Monte Carlo simulation helps deal with parameter uncertainty. In this type of simulation, parameters are sampled from their probability distributions, and as a result a greater weight is placed on more likely combinations of values. The major benefit of second order simulation is that it quantifies the total impact of parameter uncertainty on the model, thus providing a degree of confidence in the results of the analysis. Second order Monte Carlo simulation is used to perform probabilistic sensitivity analyses (see below).

1.3.5 Other Types of Microsimulation Models

Modeling of disease and interventions can also be performed using discrete-event microsimulation. In contrast to microsimulation of Markov models in which time is partitioned into intervals (cycles) during which events may or may not occur, in discrete event microsimulation the life history of individual subjects is simulated over time as they develop disease(s) using sets of equations that track the time to the next event. The equations take into consideration the demographic characteristics and risk factors of the subjects and simulate their life histories as they age, develop disease and ultimately die. Similar to Monte Carlo Markov simulation, individual events are simulated via random draws from distributions that reflect the probability of an event [29]. The impact of interventions (e.g. screening) can then be evaluated in terms of reducing disease incidence and mortality. This type of model has been used to study CRC and screening. Notable Examples include the Population Health Model (POHEM) from Canada [31] and the Microsimulation Screening Analysis (MISCAN) model from the Netherlands and the US [32].

1.4 Elements of an Economic Evaluation

There are multiple elements of an economic evaluation that are critical determinants in evaluating the quality of this type of research. These items include a clear description of 1) the study question, 2) the target population, 3) the type of economic evaluation chosen, 4) the comparators being evaluated, 5) the perspective of the analysis chosen, 6) the effectiveness data selected for analysis, 7) the time horizon of the analysis, 8) the method(s) used for evaluating health outcomes, 9) the resource and costing methodology used, 10) whether differential timing of costs and health benefits were considered (i.e. the discount rate), and 10) the manner in which variability and uncertainty was addressed. A full description of these "ingredients" can be found in the 2006 guidelines for the economic evaluation of health technologies published by the Canadian Agency for Drugs and Technologies in Health (CADTH) [33] and a brief overview is included in Appendix 2. Of note, each of these elements had to be explicitly addressed in the text of the manuscript in order for our economic evaluation of CRC screening (Chapter 3) to be considered for publication in PLoS Medicine.

1.5 Validating a Decision Analytical Model

Validating a model is a multi-stage process. The first step involves ensuring that the model structure and flow are acceptable from the viewpoint of experts in the field (i.e. face validity). The programming and formulae that comprise the model need to be carefully reviewed for syntactical errors and the results of the model should make intuitive sense. This can be an arduous task necessitating deliberate testing of the model using null or extreme input values to ensure they produce the expected outputs. This

process is commonly referred to as “debugging” the model and is intended to ensure its technical accuracy. Models should also be calibrated with the real world observations used to inform the model. Examples include national health statistics and published clinical trials. With the time horizon of the model equal to that of the cohort, the model outputs should compare favourably to the actual observations. Models that aren’t calibrated against suitable data are open to criticism. Furthermore, discrepancies between model outputs and reference data should be explained. All of the above is ultimately aimed at ensuring the internal validity of the model [34].

The ISPOR guidelines recommend that “between-model validation” be carried out when more than one model addressing a topic exists. When discrepancies are found, an effort should be made to account for them. Indeed, credibility is earned when a modeller can clarify important between-model disparities.

External validation from a modeling perspective involves comparing the model outputs to observations from external data sources that have not been used to inform the model itself. This requires the identification of suitable independent data sources, which may not always be available.

Finally, there is predictive validation. This refers to a model’s ability to predict future events. Although helpful, predictive validity is not essential since future events may depend on information that is not available at the time a model is developed and calibrated. However, the most powerful models are capable of being adapted and recalibrated as new evidence evolves.

1.6 Interpreting the Results of an Economic Evaluation

While new health care interventions brought to market usually improve health, they also tend to be offered at an increased cost to society. The figure in Appendix 3 categorizes health interventions in terms of their cost and their effect on health. This type of figure is commonly referred to as the "cost-effectiveness plane" [1]. New interventions that are more effective and less costly (quadrant II) are considered 'dominant' and should be adopted. These interventions are uncommon. Likewise, new interventions that are more expensive and less effective should be discontinued if currently used and not adopted if new. The alternatives 'dominate' such interventions (quadrant IV). In practice most interventions fall into quadrant I. That is, they increase effectiveness but also add to cost.

The incremental cost-effectiveness ratio (ICER) is the standard method for presenting the results of cost-effectiveness and cost-utility analyses. Please note that an ICER in a cost-utility analysis equates with an incremental cost per QALY. In calculating ICERs the intervention and alternatives are first rank ordered according to rising cost. An intervention's ICER is then the ratio of the difference in cost over the difference in outcome between the intervention and the next least expensive alternative. Thus, the ICER tells us how much we are paying for each additional health increment.

An ICER alone is relatively meaningless without some idea of what society considers acceptable in terms of additional expense to achieve an increment in health benefit. "An in depth discussion of this concept is beyond the scope of this thesis, but benchmarks for what constitutes an attractive ICER have been established [35]. These are in part based on what society is already willing to pay for accepted health interventions. For

example, if screening for breast cancer, a highly regarded activity among women, is associated with a cost per QALY of $10,000, would it not be worthwhile funding a new CRC screening program that costs $9000 per QALY? This question obviously becomes more complex if additional resources are unavailable, and of course, other factors are also considered within health care priority setting. At a minimum this situation demands difficult choices regarding the efficient allocation of health care dollars. In Canada, an intervention with a cost per QALY of $20,000 or less has been thought of as strongly worth considering [35].

1.7 Why Another Economic Evaluation of Colorectal Cancer Screening Was Needed.

A number of economic evaluations addressing CRC screening have been published. The National Committee on Colorectal Cancer Screening formed by Health Canada concluded in 2002 that CRC screening should be made available to Canadians 50 to 74 years of age using FOBT every 2 years. This recommendation was partly based on favourable cost-effectiveness data of a biennial FOBT screening program generated from the POHEM microsimulation model mentioned above [36]. In 2002 a systematic review for the U.S. Preventive Services Task Force concluded that each of the available CRC screening options was considered cost-effective compared to no screening. However, no one test consistently had the best ICER [37], and screening using all of the available tests were thus considered reasonable options for average risk individuals. Since that time the landscape of CRC treatment and screening has changed significantly. A number of promising new CRC screening modalities have been developed including fecal DNA and

CTC [38, 39], some tests have fallen out of favour such as barium enema [40], and important new data on the more traditional screening tests have emerged [41, 42]. Furthermore, new chemotherapeutic agents for CRC have come to market. While these treatments are more effective, they are also far more expensive, strengthening the economic argument for CRC screening.

Many of the more recently conducted economic evaluations have produced inconsistent results. Some cost-effectiveness analyses of CTC have supported use of the technology [43, 44] while others have not [45-47]. An older study looking at fecal DNA found it to be cost-effective compared to no screening, but inferior to FOBT and colonoscopy [48]. New test performance data on fecal DNA are available [49] and the test is now less expensive. Contemporary economic evaluations of flexible sigmoidoscopy are lacking in the literature given previous economic evaluations of flexible sigmoidoscopy do not incorporate the more recent data from RCTs.

There are also a number of issues with many of the recent decision analytical models which warrant discussion. Some have not considered all the commonly available CRC screening modalities [45, 50, 51], while others have ignored costs [51] and very few [52] have considered the impact of non-medical costs in the primary analysis. Furthermore, many analyses have simply inflation adjusted out dated CRC treatments rather than considering modern-day treatment costs, which is suboptimal.

A Canadian economic evaluation of CTC for CRC screening among average risk individuals [52] was recently performed to help inform decision makers in provincial jurisdictions across Canada regarding whether to fund CTC for primary CRC screening. CTC was found to be both more costly and less effective than colonoscopy [52]. We

performed the economic evaluation and budget impact analysis comprising the report. Although this major effort followed recommended guidelines for conducting economic evaluations [33], further work was needed, since that model did not include FIT, flexible sigmoidoscopy and fecal DNA, which are significant omissions. Moreover, this analysis did not properly account for the higher costs of treating advanced-stage CRC.

CHAPTER TWO: MANUSCRIPT

Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, Hemmelgarn BR. Clin Gastroenterol Hepatol. 2009 Dec;7(12):1272-8

REVIEWS

Prevalence of Adenomas and Colorectal Cancer in Average Risk Individuals: A Systematic Review and Meta-analysis

STEVEN J. HEITMAN,* PAUL E. RONKSLEY,‡ ROBERT J. HILSDEN,*,‡ BRADEN J. MANNS,*,‡ ALAA ROSTOM,* and BRENDA R. HEMMELGARN*,‡

*Departments of *Medicine and ‡Community Health Sciences, University of Calgary, Calgary, Alberta, Canada*

BACKGROUND & AIMS: There is an extensive yet inconsistent body of literature reporting on the prevalence of adenomatous polyps (adenomas) and colorectal cancer among average risk individuals. The objectives of our study were to determine the pooled prevalence of adenomas and colorectal cancer, as well as nonadvanced and advanced adenomas, among average risk North Americans. **METHODS:** Articles were obtained by searching electronic databases (MEDLINE: 1950 through March 2008 and EMBASE: 1980 through March 2008), bibliographies, major journals, and conference proceedings, with no language restrictions. Two reviewers independently selected cross-sectional studies reporting adenoma and colorectal cancer prevalence rates in average risk individuals and assessed studies for inclusion and quality, and extracted the data for analysis. Pooled adenoma and colorectal cancer prevalence rates were estimated using fixed and random effects models. Stratification and metaregression was used to assess heterogeneity. **RESULTS:** Based on 18 included studies, the pooled prevalence of adenomas, colorectal cancer, nonadvanced adenomas, and advanced adenomas was 30.2%, 0.3%, 17.7%, and 5.7%, respectively. Heterogeneity was observed in the pooled prevalence rates for overall adenomas, advanced adenomas, and colorectal cancer and was explained by the mean age ($\geq$ 65 years vs $<$ 65 years) with older cohorts reporting higher prevalence rates. None of the study quality indicators was found to be significant predictors of heterogeneity. **CONCLUSIONS: The high prevalence of advanced adenomas and colorectal cancer, especially among older screen-eligible individuals, provides impetus for expanding colorectal cancer screening programs. Furthermore, the pooled prevalence estimates can be used as quality indicators for established programs.**

Colorectal cancer (CRC) is the fourth most commonly diagnosed cancer and the second most common cause of cancer death in North America.[1,2] Screening has been shown to reduce mortality through removal of adenomatous polyps and identifying earlier stage cancers.[3–5] Furthermore, CRC screening is considered cost effective[6] and is recommended for average risk individuals 50 to 75 years old in North America.[7,8]

The prevalence of adenomas and CRC reported among average risk individuals has varied in the literature.[9,10] This may be attributed to differences in study demographics including age, gender, and possibly ethnicity. A clearer understanding of the prevalence of adenomas and CRC is important for a number of reasons. First, clinicians need to be able to inform their patients of the risk of CRC such that informed decisions regarding screening can be made. Second, a firm understanding of the prevalence of colorectal neoplasia offers an objective method for evaluating the quality of CRC screening programs. Finally, the cost effectiveness of the available screening modalities may depend on adenoma and CRC prevalence and thus provide important information for health policy decision makers.

Systematic reviews are effective tools for summarizing existing evidence. Although a systematic review of advanced adenoma incidence among individuals with a previous history of adenomatous polyps has been performed,[11] to our knowledge a systematic review among patients lacking a personal history or strong family history of adenomas and CRC has not been done. Thus, an important gap in the literature remains. The primary objective of this study, therefore, was to perform a systematic review and meta-analysis of cross sectional studies reporting the prevalence of adenomas and the prevalence of CRC among average risk North American individuals age 50 to 75. We also sought to determine the pooled prevalence of nonadvanced and advanced adenomas as a secondary objective.

Methods

Search Strategy

We performed this systematic review using a predetermined protocol and in accordance with published guidelines for reporting of observational studies.[12] We identified all potentially relevant articles regardless of publication language by searching MEDLINE including the In Process and Non-Indexed Citations (1950 through March 2008) and EMBASE (1980 through March 2008). Searches were enhanced by scanning bibliographies of identified articles and review articles, as well as reviewing conference proceedings from 2 major North American gastroenterology meetings (Digestive Disease Week and American College of Gastroenterology) and the tables of contents for 3 major gastrointestinal journals (*Gastroenterology*, *Gut*, and *American Journal of Gastroenterology*) and 2 major diagnostic

Abbreviations used in this paper: CI, confidence interval; CRC, colorectal cancer; CT, computed tomography.

1542-3565/09/$36.00
doi:10.1016/j.cgh.2009.05.032

imaging journals (*Radiology* and *American Journal of Roentgenology*) from 2005 to 2008. Experts in the field were contacted regarding missed, ongoing, or unpublished studies.

Using the strategy recommended for systematic reviews of observational studies[13] we searched the electronic databases using 3 comprehensive search themes that were then combined using the Boolean operator "and." To identify relevant testing methodology (theme 1), the first search was undertaken using the Boolean operator "or" to explode and map the Medical Subject Headings "Colonography, Computed Tomographic" or "Colonoscopy" or "Autopsy" or "Pathology" or text words "Colonoscop$" or "Colonograph$" or "CT" or "Autop$" or "Patholog$." To identify relevant outcomes (theme 2), a second search was performed using the Boolean operator "or" to explode and map the Medical Subject Headings "Polyps" or "Colonic Polyps" or "Adenomatous Polyps" or "Intestinal Polyps" or "Colonic Neoplasms" or text words "Colon$" or "Cancer$ or "Neoplasm$." To identify relevant study designs (theme 3), a final search using the Boolean operator "or" to explode and map the Medical Subject Headings "Prevalence" or "Cohort Studies" or "Cross Sectional Studies" or "Follow-up Studies" or "Incidence" or text words "Prevalence$" or "Concurrent" or "Cohort$" or "Cross Sectional" or "Survey$" or "Follow up" or "Followup" or "Incidence" or "Studies" was performed.

Study Selection

Two individuals (SJH and PER) independently reviewed identified abstracts for eligibility. All abstracts reporting original adenoma and/or CRC prevalence data among adults were selected for full text review. This initial stage was intentionally liberal; we only discarded abstracts that clearly did not meet the aforementioned criteria. The interrater agreement for this stage was high (κ = 0.80; 95% confidence interval [CI], 0.75–0.84). Disagreements were resolved by consensus.

The same reviewers then performed a full text review of articles that met the inclusion criteria and of articles for which there was uncertainty regarding eligibility. Full text articles were retained if they met the inclusion criteria of study design (cross sectional study), study population (asymptomatic average risk 50- to 75-year-old individuals), intervention (full colonic evaluation by colonoscopy or autopsy), and outcome (prevalence of adenomas and/or CRC). At least 95% of the cohort had to have undergone a colonoscopy to the cecum in order to be retained for analysis unless the article was an autopsy study. Histological confirmation of all polyps was required.

We defined average risk as asymptomatic individuals lacking high risk medical conditions (polyposis syndromes, inflammatory bowel disease, etc), and a personal or strong family history of adenomatous polyps or CRC. We excluded all studies containing >10% of individuals with a first degree relative with CRC. Authors had to specifically state that the cohort was "average risk" if details regarding family history were not provided. Autopsy studies were excluded if the premorbid CRC risk profile of the subjects was not reported. Both complete manuscripts and abstracts not yet published in full were eligible for inclusion. We attempted to contact authors of studies published only in abstract form to obtain additional study information and confirm final results.

Our primary intent was to determine adenoma and CRC prevalence estimates for asymptomatic average risk 50- to 75-year-old North Americans. We elected to exclude non-North American studies to limit potential sources of heterogeneity that could arise from differences in regional demographics (eg, diet, smoking habits).

Data Extraction and Quality Assessment

Both reviewers independently extracted data from all studies that fulfilled the inclusion criteria; disagreement was resolved by consensus. The following data were extracted: number of patients, cohort demographics, adenoma and CRC prevalence, referral type (community vs hospital-based cohort), study setting (registry, single center, or multicenter). We also abstracted details of study quality including: whether consecutive subjects were enrolled; whether explicit criteria were used to define adenoma and CRC; whether patients were referred; and whether baseline differences and important confounders were addressed.

The primary outcomes were overall adenoma prevalence and CRC prevalence. We also determined, as secondary outcomes, the prevalence of nonadvanced adenomas (<10 mm without high risk histology) and advanced adenomas (≥10 mm, villous or tubulovillous histology, or high grade dysplasia).

Data Synthesis and Analysis

The proportion of adenomas and CRC, along with 95% CI, was identified in each study. For small proportions (when the numerator is small), the calculated lower limit of a confidence interval may fall below 0 based on a Gaussian distribution. To ensure that all CIs were between 0 and 1 the Wilson score interval was calculated using a binomial distribution.[14] This has been shown to be suitable for studies with small sample size and/or extreme probabilities.[15] Weights for the individual studies were calculated using the inverse of the variance method. To obtain a pooled estimate of adenoma and CRC prevalence, a fixed effect model was initially performed. The Q statistic was calculated to assess for significant heterogeneity between the included studies. In light of the heterogeneity observed, we also used a random effects model to obtain a pooled estimate of prevalence when appropriate.

We performed metaregression to explore clinical and methodological factors contributing to heterogeneity. Factors assessed included consecutive versus nonconsecutive patient referral, single versus multicenter recruitment, sample size ($n \geq 100$ vs $n < 100$), and cohort demographics (age and gender). We also assessed the potential importance of study quality factors including the extent to which the study population was described and whether or not potential confounders were addressed.

All analyses were performed using STATA 10.0 (Statacorp, College Station, TX). $P \leq .05$ was considered statistically significant.

Results

Identification of Studies

The progress through stages of the systematic review is summarized in Figure 1. Our initial search yielded a total of 4710 citations. After an initial screen, 333 met criteria for full text review, of which 298 were excluded. None of the autopsy studies could be included given that the premorbid state of the included subjects was not reported. Three studies based on the Clinical Outcomes Research Initiative (CORI) electronic regis-

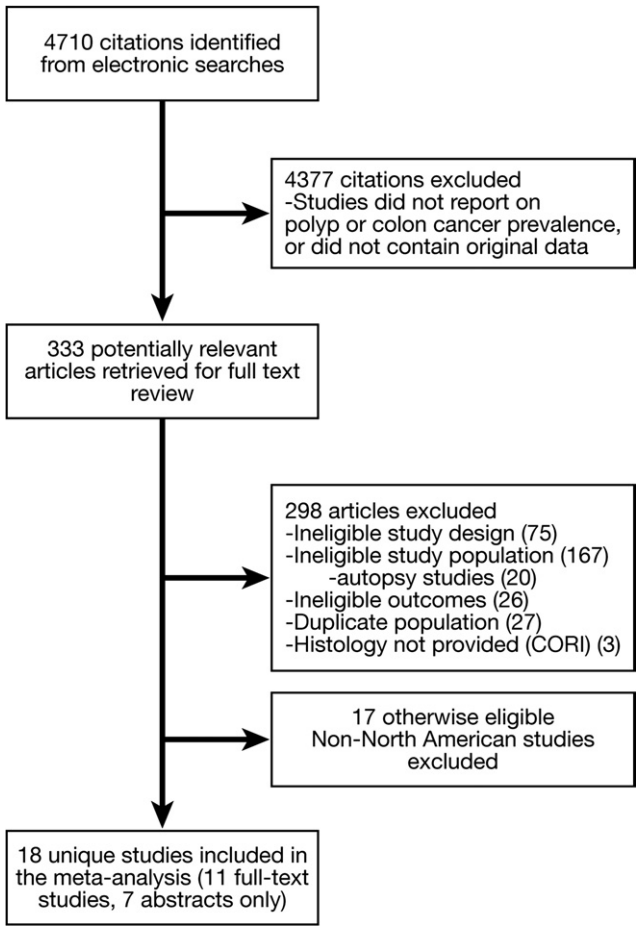


Figure 1. Article flow diagram.

try[16–18] were identified and were not included in the final analysis because histopathology was not available. When duplicate reporting of the same cohort of patients was apparent, we selected the largest study containing the most information on unique subjects. Seventeen studies meeting all of the inclusion criteria were from countries outside of North America. As such, 18 studies reporting on adenoma and/or CRC prevalence among North American populations were used for the present analysis.[10,19–35] Of these, 11 were available in full manuscript[10,20,22,24–28,30,31,33] and 7 were published in abstract form only.[19,21,23,29,32,34,35] No unpublished or ongoing studies were identified.

Details of Included Studies

Table 1 shows details of the studies that met our inclusion criteria. The publication dates ranged from 1990 to 2008, with the number of patients per study ranging from 84 to 4404. All 18 studies were based in the United States. The proportion of male subjects within each study cohort ranged from 40.5% to 100%. Only 8 of the 18 studies reported a mean age for the subjects (range 56.9 to 68.6 years). Although we had initially intended to exclude results for patients over age 75, none of the included studies stratified by age categories to permit this. No studies in our initial search and none of the results on the included studies were lost due to an upper age cutoff.

Not all studies reported data on all outcomes of interest. As a result, different combinations of studies were pooled to produce our prevalence estimates for polyps (nonadvanced and advanced) and CRC. Of the 18 studies, 14 reported on overall adenoma prevalence; CRC prevalence was reported in 17 of the studies. Nonadvanced adenomas and advanced adenomas were reported in 6 and 12 of the 18 studies, respectively.

Study Quality Assessment

The quality of the studies varied (Table 2). Seven of the 18 studies did not report whether patient recruitment was consecutive. Only 6 of the 18 studies provided demographic details beyond the age and sex of the cohorts. Furthermore, age and gender-specific prevalence values were not reported in the majority of studies. Therefore, we examined differences between older and younger cohorts (age $<$ 65 vs $\geq$ 65), and male-dominated (>2/3 males) studies. Finally, other potential confounders such as race, body mass index, and smoking status were not addressed in over half of the studies.

Prevalence of Adenomas and CRC

The prevalence of adenomas among the 14 studies ranged from 22.2% to 58.2% (Table 1). When individual studies were combined in a meta-analysis there was significant heterogeneity as defined by the Q statistic (Q statistic: $P < .001$). The pooled estimate of adenoma prevalence using a random effects model was 30.2% (95% CI, 27.1%–33.3%; Figure 2*A*). Colon cancer prevalence among the 17 studies ranged from 0% to 1.7%. When combined in a meta-analysis, there was significant heterogeneity (Q statistic: $P = .012$) and therefore a pooled estimate of 0.3% (95% CI, 0.2%–0.5%) was calculated using a random effects model (Figure 2*B*). A sensitivity analysis was performed in which the prevalence estimates were determined separately for full text manuscripts and abstracts, with similar results obtained, therefore combined results only are presented here.

Prevalence of Nonadvanced and Advanced Adenomas

Nonadvanced adenoma prevalence ranged from 16.3% to 19.9% among the 6 studies. No heterogeneity was observed when these values were combined in a meta-analysis (Q statistic: $P = .465$). Using a fixed effects model, we calculated a pooled estimate of 17.7% (95% CI, 16.7%–18.6%; Figure 2*C*). Advanced adenoma prevalence ranged from 2.5% to 9.7% in 12 studies. When these studies were combined in a meta-analysis, there was significant heterogeneity (Q statistic: $P < .001$). The pooled estimate using a random effects model was 5.7% (95% CI, 4.1%–7.4%; Figure 2*D*). A sensitivity analysis was performed in which the prevalence estimates were determined separately for full text manuscripts and abstracts, with similar results obtained, therefore combined results only are presented here.

Methodological and Clinical Sources of Heterogeneity

A number of factors were assessed using metaregression to explain potential sources of underlying heterogeneity in our pooled estimates of adenoma, CRC, and advanced adenoma prevalence. None of the methodological factors assessed was identified as significant sources of heterogeneity; however, stud-

Table 1. Characteristics of Included Studies

Study	Demographics of study subjects			Outcomes reported			
	Number of patients	Mean age (*y*)	Male (%)	Adenoma prevalence (%)	CRC prevalence (%)	Nonadvanced adenoma prevalence (%)	Advanced adenoma prevalence (%)
Pickhardt et al (2004)[27]	1201	57.8	59.0	NR	0.162	NR	3.89
Mehran et al (2003)[26]	91	NR	49.5	58.2	1.10	NR	3.30
DiSario et al (1991)[20]	119	NR	100	41.2	1.68	NR	NR
Lieberman et al (1991)[25]	84	64.0	100	41.7	NR	NR	NR
Imperiale et al (2004)[10]	4404	68.6	44.6	27.0	0.704	17.3	9.67
Johnson et al (1990)[22]	90	65.0	67.8	22.2	1.10	18.9	8.89
Prajapati et al (2003)[28]	257	62.0	40.5	23.7	0.389	17.1	6.23
Rex et al (2000)[30]	121	60.5	43.0	34.7	0.00	NR	2.48
Stevens et al (2003)[33]	272	NR	49.8	24.6	0.368	18.8	5.88
Rex et al (1993)[31]	496	NR	62.5	25.2	0.202	16.3	8.87
Kim et al (2007)[24]	3163	58.1	44.4	NR	0.126	NR	3.70
Spellman et al (2007)[32,a]	3968	NR	46	22.0	0.479	NR	NR
Kang et al (2005)[23,a]	809	NR	NR	NR	0.247	NR	NR
Winston et al (2005)[35,a]	1000	NR	43	28.2	0.700	19.9	8.30
Herlihy et al (2005)[21,a]	792	NR	50.6	33.1	0.379	NR	2.65
Randall et al (2005)[29,a]	1324	NR	NR	24.2	0.227	NR	NR
Anderson et al (2008)[19,a]	600	56.9	41.8	37.8	0.00	NR	NR
Wehbi et al (2006)[34,a]	2547	NR	NR	NR	0.628	NR	6.20

NR, not reported.
[a]Abstract only.

ies with less than 100 subjects tended to report higher prevalence rates (Table 3). Of the clinical factors, age was a significant source of heterogeneity for all 3 outcomes ($P < .001$). Cohorts with a mean age of ≥65 years reported higher prevalence rates compared with studies with a mean age of <65 years for advanced adenomas (8.2% vs 3.8%) and CRC (0.7% vs 0.1%). This relationship was not found for overall adenomas (Table 3), which demonstrated a higher prevalence in the younger group. These results however were driven by the study by Mehran et al,[26] which reported on a younger cohort that had a higher prevalence of adenomas. When Mehran et al was excluded age was no longer a significant source of heterogeneity for overall adenomas. Gender was not a significant source of heterogeneity; however, studies that were composed predominantly of male subjects (>2/3 male) also tended to reported higher prevalence rates (Table 3).

Table 2. Cohort Details and Quality Indicators of Included Studies

Study	Cohort details		Study quality indicators						
	Referral type	Study setting	Sample described	Consecutive patients	Referred patients	Demographics described	Polyp histology reported	Baseline differences	Confounders discussed
Pickhardt et al (2004)[27]	Community	Multi center	Yes	Yes	Yes	Yes	Yes	Yes	No
Mehran et al (2003)[26]	Hospital	Single center	Yes	NR	NR	No	Yes	No	No
DiSario et al (1991)[20]	Community	Single center	Yes	NR	Yes	No	Yes	No	No
Lieberman et al (1991)[25]	Community	Multi center	Yes	NR	Yes	Yes	Yes	Yes	Yes
Imperiale et al (2004)[10]	Community	Multi center	Yes	NR	Yes	Yes	Yes	Yes	Yes
Johnson et al (1990)[22]	Community	Single center	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prajapati et al (2003)[28]	Community	Single center	Yes	Yes	NR	No	Yes	No	No
Rex et al (2000)[30]	Community	Multi center	Yes	No	Yes	No	Yes	No	No
Stevens et al (2003)[33]	Hospital	Single center	Yes	NR	NR	No	Yes	No	No
Rex et al (1993)[31]	Community	Single center	Yes	NR	No	Yes	Yes	Yes	Yes
Kim et al (2007)[24]	Community	Single center	Yes	Yes	Yes	No	Yes	No	No
Spellman et al (2007)[32,a]	Community	Single center	No	Yes	Yes	No	Yes	No	No
Kang et al (2005)[23,a]	Community	Single center	No	Yes	Yes	No	Yes	No	No
Winston et al (2005)[35,a]	Community	Single center	Yes	Yes	Yes	No	Yes	No	No
Herlihy et al (2005)[21,a]	Community	Single center	Yes	Yes	Yes	No	Yes	No	No
Randall et al (2005)[29,a]	Community	Single center	No	Yes	Yes	No	No	No	No
Anderson et al (2008)[19,a]	Community	Single center	Yes	NR	Yes	Yes	Yes	Yes	Yes
Wehbi et al (2006)[34,a]	Community	Multi center	No	Yes	NR	No	Yes	No	Yes

NR, not reported.
[a]Abstract only.

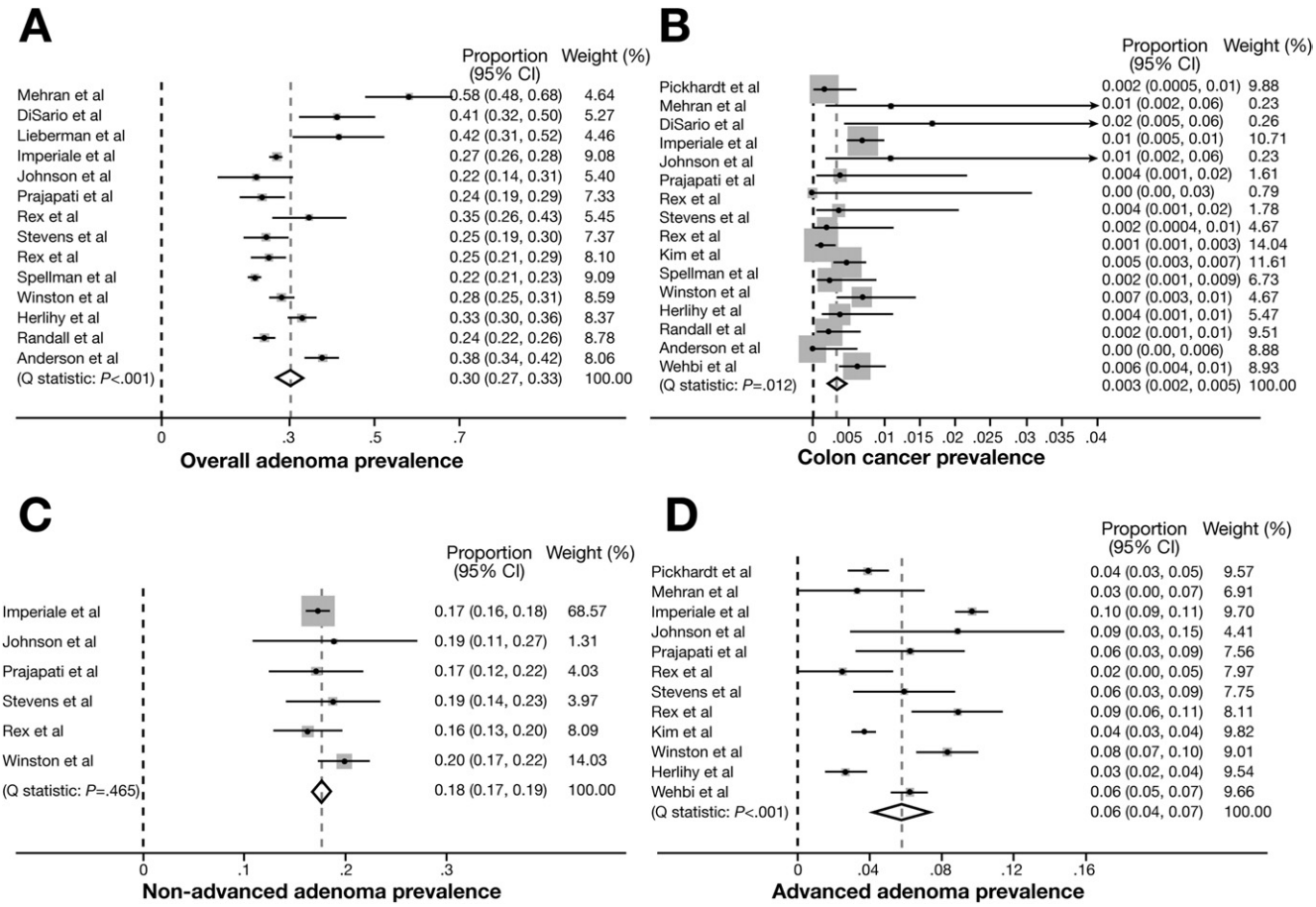


Figure 2. Forest plots of (*A*) Overall adenoma prevalence (random effects model); (*B*) Colon cancer prevalence (random effects model); (*C*) nonadvanced adenoma prevalence (fixed effects model); and (*D*) advanced adenoma prevalence (random effects model). Please note that the sum of nonadvanced and advanced adenoma prevalence does not equal overall adenoma prevalence as different combinations of studies were pooled for each estimate.

Discussion

In this, the first systematic review and meta-analysis of North American subjects at average risk for CRC, we determined pooled prevalence rates of adenomatous polyps and CRC. The overall prevalence of adenomas was 30.2% based on 14 studies reporting on 13,618 individuals and the prevalence of CRC was 0.3% based on 17 studies reporting on 21,254 individuals. The prevalence of nonadvanced and advanced adenomas was 17.7% based on 6 studies reporting on 6519 individuals and 5.7% based on 12 studies reporting on 14,434 individuals, respectively.

A number of steps were taken to ensure our systematic review was limited to average risk individuals. We excluded autopsy studies as it was not possible to determine whether the study subjects were truly average risk. Also by excluding studies with more than 10% of individuals having a first degree relative

Table 3. Stratified Meta-Analysis of Adenoma and Colon Cancer Prevalence By Demographic and Methodological Sources[a]

Cohort stratifications	Overall adenoma, % (95% CI)	CRC, % (95% CI)	Nonadvanced adenoma, % (95% CI)	Advanced adenoma, % (95% CI)
≥2/3 male	34.9 (21.8–47.9)	1.4 (0.2–2.9)	18.9 (10.8–27.0)	8.9 (3.0–14.8)
<2/3 male	30.1 (26.5–33.7)	0.3 (0.2–0.4)	17.7 (16.6–18.9)	5.5 (3.6–7.5)
≥65 years	26.6 (23.8–29.3)[b]	0.7 (0.4–0.9)[b]	17.3 (16.2–18.4)	8.2 (5.4–11.1)[b]
<65 years	34.0 (25.7–42.4)[b]	0.1 (0.0–0.2)[b]	17.1 (12.5–21.7)	3.8 (3.1–4.5)[b]
Large sample size (≥100)	28.5 (25.6–31.4)	0.3 (0.2–0.5)	17.7 (16.6–18.9)	5.8 (4.0–7.5)
Small sample size (<100)	40.6 (19.3–61.8)	1.1 (0.4–2.6)	18.9 (10.8–27.0)	5.6 (0.2–11.0)

NOTE. The sum of nonadvanced and advanced adenoma prevalence does not equal overall adenoma prevalence as different combinations of studies were pooled for each estimate.

[a]Based on random effects modeling.

[b]Significant sources of heterogeneity identified by metaregression ($P < .05$).

with CRC, we are confident that our pooled estimates are based on a relatively "pure" average risk sample.

Although current North American guidelines recommend choice among available screening modalities,[7,8] we limited our systematic review to studies in which a complete colonoscopy was used for screening. However, colonoscopy is itself an imperfect test. Studies that report the combined findings of computed tomography (CT) colonography and colonoscopy can provide a more complete assessment of the entire colon by reducing the false negative rate of colonoscopy.[36] For this reason we included CT colonography-related Medical Subject Heading terms in order to ensure that these studies were identified in our initial search. Studies where colonoscopy was only performed as a follow-up procedure (ie, following a positive fecal occult blood test or when a polyp was found on sigmoidoscopy or suspected on barium enema or CT colonography) were excluded given the possibility for higher prevalence estimates among this higher risk group. Including studies where only flexible sigmoidoscopy was used could have underestimated our prevalence estimates given that 2%–5% of patients may have isolated proximal neoplasia beyond the reach of a sigmoidoscope.[37]

In this meta-analysis, we found significant heterogeneity among the 18 studies reporting on the prevalence of adenomas, advanced adenomas, and CRC. Male gender is a known risk factor for adenomatous polyps.[38] Unfortunately, the included studies did not report data for males and females separately and thus we could not stratify by gender. Most studies did however report the proportion of males and females. Although it did not reach statistical significance, when studies with $>2/3$ males were compared with studies that had $<2/3$ male participants, the rates for overall adenomas, advanced adenomas, and CRC were higher among male-dominated studies (Table 3).

The mean age of the cohort was identified to be an important source of heterogeneity for these prevalence estimates. Age is perhaps the single most important independent determinant of adenoma prevalence.[39] This was particularly true for the prevalence of advanced adenomas and CRC in which the rates were 2- and 7-fold higher among older cohorts, respectively (Table 3).

Our study has limitations. First, the reporting of prevalence by type of neoplasia (ie, adenomas and/or CRC) was not complete across all studies. In fact, only 6 of the 18 articles reported all the outcomes of interest. Whereas data on overall adenoma, advanced adenoma, and CRC prevalence were either reported or could be extracted in the majority of studies, only 6 of the 18 studies reported on nonadvanced adenoma prevalence separately. Thus, different combinations of studies had to be pooled to calculate each of our prevalence estimates. This explains why the sum of nonadvanced and advanced adenoma prevalence does not equate to the overall adenoma prevalence that we report. Furthermore, our pooled nonadvanced adenoma prevalence may not be accurate given the limited data that were available for analysis. Definitions for advanced adenomas were similar across studies. However, some reported on advanced neoplasia, which included invasive carcinomas. We took this into consideration by excluding these values from the advanced adenoma proportions. Second, many of the factors that we attempted to extract from studies were either not available or not explicitly stated by the authors. As such, we could not reliably test for these potentially important study quality indicators in our metaregression. Third, the findings of cross sectional studies are subject to both known and unknown confounders. Age was identified as an important confounder. However, we were limited in our ability to stratify by narrower age categories and by gender. In addition, we were unable to explore other potential confounders such as race, body mass index, smoking status, and medication use including nonsteroidal anti-inflammatory drugs, as the data were not generally provided. Finally, our metaregression was likely underpowered and unable to detect the significance of some of the moderate sources of heterogeneity given that the number of pooled studies in each analysis was relatively small.

In addition to summarizing this large body of literature and accounting for sources of heterogeneity, this systematic review and meta-analysis is important for a number of reasons. First, our data provide important information for practitioners to educate the public regarding the importance of CRC screening. Adherence to CRC screening recommendations remains low.[40] This may be particularly relevant for those over age 65 given that our results suggest that 1 in every 150 average risk patients screened would have prevalent CRC and 1 in every 12 patients harbor advanced adenomas, lesions known to be at significant risk of progression to CRC. Secondly, our data provide important baseline inputs for decision analytic models evaluating the cost effectiveness of current and emerging CRC screening modalities. Finally, population-based CRC screening programs are becoming increasingly common. Our pooled prevalence estimates offer an important performance metric for evaluating new and existing programs. Although all studies were conducted in the United States, we believe that our results are generalizable across North America. Programs reporting lower adenomatous polyp and CRC prevalence among average risk individuals need to explore reasons for their lower detection rates.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2007. CA Cancer J Clin 2007;57:43–66.
2. Canadian Cancer Society, National Cancer Institute of Canada. Canadian Cancer Statistics 2007. Toronto, Canada: Canadian Cancer Society, 2007.
3. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472–1477.
4. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467–1471.
5. 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;328:1365–1371.
6. Pignone M, Saha S, Hoerger T, et al. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:96–104.
7. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale—update based on new evidence. Gastroenterology 2003;124:544–560.
8. Leddin D, Hunt R, Champion M, et al. Canadian Association of Gastroenterology and the Canadian Digestive Health Foundation: Guidelines on colon cancer screening. Can J Gastroenterol 2004; 18:93–99.

9. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003;349:2191–2200.
10. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;351:2704–2714.
11. Saini SD, Kim HM, Schoenfeld P. Incidence of advanced adenomas at surveillance colonoscopy in patients with a personal history of colon adenomas: a meta-analysis and systematic review. Gastrointest Endosc 2006;64:614–626.
12. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008–2012.
13. Egger M, Smith GD, Altman DG. Systematic reviews in health care: meta-analysis in context. London: BMJ Publishing Group; 2001.
14. Wilson E. Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 1927;22:209–212.
15. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. The American Statistician 1998;52:119–126.
16. Lieberman DA, de Garmo PL, Fleischer DE, et al. Colonic neoplasia in patients with nonspecific GI symptoms. Gastrointest Endosc 2000;51:647–651.
17. Lieberman DA, Holub J, Eisen G, et al. Prevalence of polyps greater than 9 mm in a consortium of diverse clinical practice settings in the United States. Clin Gastroenterol Hepatol 2005; 3:798–805.
18. Thornton JG, Morris AM, Thornton JD, et al. Racial variation in colorectal polyp and tumor location. J Natl Med Assoc 2007;99: 723–728.
19. Anderson JC, Stein B, Rajapakse R, et al. BMI and advanced colorectal neoplasia in average risk women: results of screening colonoscopy in a diverse population using a high resolution colonoscope. Gastroenterology 2008;134:A160.
20. DiSario JA, Foutch PG, Mai HD, et al. Prevalence and malignant potential of colorectal polyps in asymptomatic, average-risk men. Am J Gastroenterol 1991;86:941–945.
21. Herlihy KJ, Yarze JC, Fritz HP, et al. Colonoscopic screening of a community-based population of asymptomatic average-risk individuals. Am J Gastroenterol 2005;100:1066.
22. Johnson DA, Gurney MS, Volpe RJ, et al. A prospective study of the prevalence of colonic neoplasms in asymptomatic patients with an age-related risk. Am J Gastroenterol 1990;85:969–974.
23. Kang HS, Cho H, Abdrabbo K. The incidence of adenomatous colon polyps in average risk patients referred for colorectal cancer screening: Experience of a single endoscopist at a university center. Am J Gastroenterol 2005;100:382.
24. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007;357:1403–1412.
25. Lieberman DA, Smith FW. Screening for colon malignancy with colonoscopy. Am J Gastroenterol 1991;86:946–951.
26. Mehran A, Jaffe P, Efron J, et al. Screening colonoscopy in the asymptomatic 50- to 59-year-old population. Surg Endosc 2003; 17:1974–1977.
27. Pickhardt PJ, Choi JR, Hwang I, et al. Nonadenomatous polyps at CT colonography: prevalence, size distribution, and detection rates. Radiology 2004;232:784–790.
28. Prajapati DN, Saeian K, Binion DG, et al. Volume and yield of screening colonoscopy at a tertiary medical center after change in medicare reimbursement. Am J Gastroenterol 2003;98:194–199.
29. Randall CW, Taboada CM, Garza C, et al. Histology, distribution and incidence of polyps in average risk patients undergoing colorectal cancer screening. Am J Gastroenterol 2005;100:1074.
30. Rex DK, Khan AM, Shah P, et al. Screening colonoscopy in asymptomatic average-risk African Americans. Gastrointest Endosc 2000;51:524–527.
31. Rex DK, Lehman GA, Ulbright TM, et al. Colonic neoplasia in asymptomatic persons with negative fecal occult blood tests: influence of age, gender, and family history. Am J Gastroenterol 1993;88:825–831.
32. Spellman SJ, Bader M, Zogg DI. Outcomes and complications in average risk colon cancer screning in a community hospital. Am J Gastroenterol 2007;102:1194.
33. Stevens T, Burke CA. Colonoscopy screening in the elderly: when to stop? Am J Gastroenterol 2003;98:1881–1885.
34. Wehbi MA, Nam JW, Nesmith CG, et al. The utilization of colonoscopy as a screening method for colorectal cancer in two U.S. metropolitan teaching hospitals. Gastroenterology 2006;130: A616.
35. Winston DH, Baughman C, Siddoway R, et al. Colonoscopy screening of average-risk patients for colorectal cancer in a large staff model HMO. Am J Gastroenterol 2005;100:1061.
36. Pineau BC, Paskett ED, Chen GJ, et al. Validation of virtual colonoscopy in the detection of colorectal polyps and masses: rationale for proper study design. Int J Gastrointest Cancer 2001; 30:133–140.
37. Lewis JD, Ng K, Hung KE, et al. Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy. Arch Intern Med 2003;163:413–420.
38. Rex DK. Colonoscopy: a review of its yield for cancers and adenomas by indication. Am J Gastroenterol 1995;90:353–365.
39. Itzkowitz SH. Colonic polyps and polyposis syndromes. 7th ed. Philadelphia: Saunders, 2002.
40. Rabeneck L, Paszat LF. A population-based estimate of the extent of colorectal cancer screening in Ontario. Am J Gastroenterol 2004;99:1141–1144.

Reprint requests

Address requests for reprints to: Brenda Hemmelgarn, MD, PhD, FRCPC, Department of Medicine, 1403 29th Street NW, Calgary, Alberta, Canada T2N 2T9. e-mail: brenda.hemmelgarn@albertahealthservices.ca; fax: (403) 944-2876.

Acknowledgments

Steven Heitman and Paul Ronksley contributed equally to this article.

Conflicts of interest

The authors disclose no conflicts.

Funding

Steven Heitman was supported by a Clinical Fellowship Award from the Alberta Heritage Foundation for Medical Research, and Paul Ronksley by an Achievers in Medical Science Graduate Scholarship. Brenda Hemmelgarn and Robert Hilsden are supported by Population Health Investigator Awards from the Alberta Heritage Foundation for Medical Research, and Brenda Hemmelgarn and Braden Manns are supported by New Investigator Awards from the Canadian Institutes of Health Research.

CHAPTER THREE: MANUSCRIPT

Colorectal cancer screening for average-risk North Americans: an economic evaluation. Heitman SJ, Hilsden RJ, Au F, Dowden S, Manns BJ. PLoS Med 2010 Nov 23;7(11):e1000370.

Colorectal Cancer Screening for Average-Risk North Americans: An Economic Evaluation

Steven J. Heitman[1,2], Robert J. Hilsden[1,2], Flora Au[1], Scot Dowden[1,3], Braden J. Manns[1,2,4]*

1 The Department of Medicine, University of Calgary, Alberta, Canada, **2** The Department of Community Health Sciences, University of Calgary, Alberta, Canada, **3** Alberta Health Services - Cancer Care, Alberta, Canada, **4** Libin Cardiovascular Institute, University of Calgary, Alberta, Canada

Abstract

Background: Colorectal cancer (CRC) fulfills the World Health Organization criteria for mass screening, but screening uptake is low in most countries. CRC screening is resource intensive, and it is unclear if an optimal strategy exists. The objective of this study was to perform an economic evaluation of CRC screening in average risk North American individuals considering all relevant screening modalities and current CRC treatment costs.

Methods and Findings: An incremental cost-utility analysis using a Markov model was performed comparing guaiac-based fecal occult blood test (FOBT) or fecal immunochemical test (FIT) annually, fecal DNA every 3 years, flexible sigmoidoscopy or computed tomographic colonography every 5 years, and colonoscopy every 10 years. All strategies were also compared to a no screening natural history arm. Given that different FIT assays and collection methods have been previously tested, three distinct FIT testing strategies were considered, on the basis of studies that have reported "low," "mid," and "high" test performance characteristics for detecting adenomas and CRC. Adenoma and CRC prevalence rates were based on a recent systematic review whereas screening adherence, test performance, and CRC treatment costs were based on publicly available data. The outcome measures included lifetime costs, number of cancers, cancer-related deaths, quality-adjusted life-years gained, and incremental cost-utility ratios. Sensitivity and scenario analyses were performed. Annual FIT, assuming mid-range testing characteristics, was more effective and less costly compared to all strategies (including no screening) except FIT-high. Among the lifetimes of 100,000 average-risk patients, the number of cancers could be reduced from 4,857 to 1,782 and the number of CRC deaths from 1,393 to 457, while saving CAN$68 per person. Although screening patients with FIT became more expensive than a strategy of no screening when the test performance of FIT was reduced, or the cost of managing CRC was lowered (e.g., for jurisdictions that do not fund expensive biologic chemotherapeutic regimens), CRC screening with FIT remained economically attractive.

Conclusions: CRC screening with FIT reduces the risk of CRC and CRC-related deaths, and lowers health care costs in comparison to no screening and to other existing screening strategies. Health policy decision makers should consider prioritizing funding for CRC screening using FIT.

Please see later in the article for the Editors' Summary.

Citation: Heitman SJ, Hilsden RJ, Au F, Dowden S, Manns BJ (2010) Colorectal Cancer Screening for Average-Risk North Americans: An Economic Evaluation. PLoS Med 7(11): e1000370. doi:10.1371/journal.pmed.1000370

Academic Editor: Eduardo L. Franco, McGill University, Canada

Received June 3, 2010; **Accepted** October 14, 2010; **Published** November 23, 2010

Funding: This project was supported by the Canadian Agency for Drugs and Technologies in Health and Alberta Innovates - Health Solutions. BJM and RJH are supported by Alberta Innovates - Health Scholar Awards. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: CRC, colorectal cancer; CTC, computed tomographic colonography; FDNA, fecal DNA assay; FOBT, fecal occult blood test; FIT, fecal immunochemical test; QALY, quality-adjusted life-year; RCT, randomized controlled trial

* E-mail: braden.manns@albertahealthservices.ca

Introduction

As the fourth most common cancer and second-leading cause of cancer death among men and women [1], colorectal cancer (CRC) is an important health issue. CRC fulfills the World Health Organization (WHO) criteria for mass screening [2], and existing clinical practice guidelines recommend that average risk individuals begin screening at age 50 [3–6]. A variety of CRC screening modalities are available, including stool-based tests and radiological and endoscopic examinations of the colon. Colonoscopy has high sensitivity for identifying adenomas and cancer and permits the removal of polyps during a screening examination [7]. However, the risk of complications (including bleeding, perforation, and death) and barriers to access including limited availability and high patient-borne costs [8] diminish its appeal. The guaiac-based fecal occult blood tests (FOBTs) have been shown in randomized controlled trials (RCTs) to reduce CRC mortality [9–11]. However, FOBT has low sensitivity for identifying colorectal neoplasia, in particular adenomas. The fecal immunochemical tests (FITs) have improved test performance characteristics [12] and potential to improve participation rates compared to FOBT and flexible sigmoidoscopy [13]. A third type of stool test, based on the detection of DNA shed by neoplastic tissue (fecal DNA) is also available [14,15]. Lastly, computed tomographic colonography (CTC) or "virtual" colonoscopy is a promising new modality [6]. Although recent studies [16–18] have shown CTC to rival colonoscopy in detecting advanced adenomas and CRC, CTC is expensive, requires a full colonic preparation, and the available cost-effectiveness data have been contradictory [19–21].

In light of the rapidly rising costs of chemotherapy for CRC [22], and evidence that CRC mortality can be reduced by screening [9–11], population-based screening programs for average risk individuals are being considered in several countries. In the absence of firm comparative evidence to guide the selection of any one modality, the practice in some jurisdictions has been to recommend choice among the available screening options [3–5]. However, some countries do not support population-based CRC screening and many with organized programs do not offer choice [23]. Given the varied test performance characteristics and the significant differences in costs and resources associated with each, health care decision makers should consider the results of cost-effectiveness analyses when deciding whether or not to offer screening and in selecting the most appropriate screening modality.

There have been several previous economic analyses of CRC screening [24], though recent studies have failed to consider all potentially relevant strategies including CTC [25,26] and FIT [27]. Furthermore, a wide range of FIT test performance has been reported, the impact of which requires further exploration in cost-effectiveness analyses. Finally, many studies have not considered current CRC treatment costs, nor the different nonmedical costs between CRC screening strategies, both of which may be important. Given these limitations, we performed a full economic evaluation of all relevant CRC screening modalities in North America, and present our results in a transparent fashion to assist medical decision makers.

Methods

Overview

An incremental cost-utility analysis was performed comparing the following CRC screening modalities: guaiac-based FOBT, FIT, fecal DNA, colonoscopy, flexible sigmoidoscopy, and CTC. These modalities were compared to each other and to a no screening natural history arm among average-risk individuals, aged 50 to 75 y. Two average-risk age-stratified patient cohorts were simultaneously modelled: people aged 50–64 and 65–75. In the base case, screening was assumed to continue from age 50–75, but the analysis continued over the lifetime of the cohorts. Average risk was defined as asymptomatic individuals with no personal or family history of CRC or adenomatous polyps and no history of preexisting medical conditions known to increase the risk of CRC (e.g., inflammatory bowel disease).

Although we acknowledge that many jurisdictions are already committed to CRC screening, we included a no screening strategy given that, despite widespread screening recommendations, the majority of individuals are not being screened [28]. In the base case analysis, costs were those relevant to a publicly funded health care system and included patient time and travel costs in keeping with recent guidelines [29]. Consistent with contemporary guidelines and the perspective of the publicly funded health care system, costs resulting from lost productivity were not considered [29]. Given the impact of CRC on both quantity and quality of life, health benefits were measured in quality-adjusted life-years (QALYs) gained over a lifetime horizon. Future costs and benefits were discounted at 5% annually [29]. Base case analyses were performed using Markov cohort simulation; second order probabilistic sensitivity analysis was used to derive 95% confidence intervals around mean costs and QALYs, and for probabilistic sensitivity analysis (see below). First order Monte Carlo simulation was used to estimate CRC incidence and death rates and the number of primary screening tests and colonoscopies required. Incremental analyses (expressed as the cost per QALY gained) were performed by rank ordering all competing strategies by increasing cost after eliminating strategies that were more costly and less effective (i.e., dominated).

Model Validation

Consistent with guidelines for good modeling in health care [30], the validity of our model was formally established including extensive "debugging" exercises and calibration to published clinical datasets [9–11]. Gastroenterologists, including two of the authors (SJH and RJH), carefully reviewed the structure and flow of the model. The model was also reviewed by Alaa Rostom, Gastroenterologist and Medical Director at the Forzani and MacPhail Colon Cancer Screening Centre in Calgary, Alberta. Ultimately, it was determined that the model had good face validity. After ensuring that there were no syntactical errors, we first calibrated the model's no screening arm against the no screening control arms of the landmark FOBT RCTs [9–11]. For this we used baseline adenoma and CRC prevalence rates from a contemporary meta-analysis [31] and ensured that the number of cancers and cancer deaths generated by our model closely approximated the control arms of the clinical trials over an identical follow-up period. We next ensured that the number of cancers and cancer deaths predicted by the FOBT screening arms closely approximated those noted within the FOBT arms of the FOBT RCTs. All of the other strategies were validated in a similar fashion assuring face validity and calibration. Finally, we also compared our CRC and CRC death rate with those generated by another validated decision analytical model, noting near perfect correlation [32].

Computer Simulation Model

The Markov model was constructed using decision analysis software (TreeAge Pro Suite 2007). It was assumed that all CRCs arise through the following sequence: normal colon → nonadvanced

adenoma → advanced adenoma → CRC. Nonadvanced adenomas were defined as tubular adenomas <10 mm in size. Advanced adenomas comprised any adenoma ≥10 mm regardless of histology, and adenomas <10 mm containing at least 25% villous component and/or high grade dysplasia. We considered several general health states, including (1) alive with no prevalent or prior history of adenomas or CRC, (2) alive with a missed adenoma, (3) alive with a missed asymptomatic CRC, (4) alive with a missed CRC after presenting with symptoms, (5) alive with a CRC found by screening, (6) alive post polypectomy, and (7) dead. Each year (1-y cycle length), individuals with or without adenomas or CRC could either remain in the same health state, progress to another health state, or die (Figure 1).

In the base case, screening was offered annually for FOBT and FIT, every 3 y for fecal DNA, every 5 y for flexible sigmoidoscopy and CTC, and every 10 y for colonoscopy. Once a patient was diagnosed with either an adenoma or CRC, the model's design permitted subsequent surveillance with colonoscopy at either 3- or 5-y intervals depending on the results of the last colonoscopy, consistent with current guidelines [4–6]. Screening and surveillance commenced at age 50 and stopped at age 75.

Data Inputs

Risk of polyps and CRC and the adenoma-carcinoma sequence. We based our prevalence estimates of adenomatous polyps and CRC on a recent systematic review among those at average risk for CRC [31]. Age was determined to be an important source of heterogeneity in the pooled estimates [31], and thus the prevalence rates in our model were stratified into two age categories: 50–64 and 65–75 y (Table 1).

Not all polyps are adenomatous. However, determining a polyp's histology generally requires that it be biopsied or removed. As a result, some polypectomies expose patients to complications without reducing the risk of CRC. We estimated that 41% [33] of polyps <10 mm were adenomatous compared to 82% of polyps ≥10 mm (Table 1) [16]. Screening guidelines recommend that all polyps be removed at the time of a colonoscopy to determine histology and establish an appropriate surveillance interval. Although some advocate for ignoring polyps <5 mm in size found on CTC, we assumed that all patients with polyps found on CTC regardless of size would be referred for colonoscopy. The risk of proximal adenomatous polyps and CRC is increased among those with adenomatous polyps in the left colon [34]. As such, we assumed that patients with left-sided adenomas found on flexible sigmoidoscopy would be referred for colonoscopy consistent with general clinical practice.

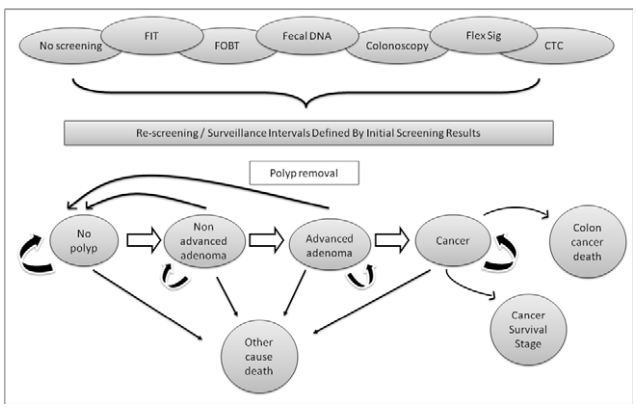


Figure 1. Model bubble diagram. This diagram depicts the general health states and flow through the model.
doi:10.1371/journal.pmed.1000370.g001

The rate of progression of adenomatous polyps is not well established. We initially chose progression rates that were consistent with other published models [32], and made small adjustments to these rates to ensure that the total number of CRCs in our natural history/no screening strategy closely approximated the number of CRCs found in the control arms of the FOBT trials [27].

Mortality. Death occurred according to either age-dependent population mortality rates observed for Canadians [35] or based on the mortality rates observed for patients with CRC according to their stage at diagnosis (Table 1) [36]. Those with CRC found through screening were assumed to have improved survival over patients presenting with symptomatic cancer, on the basis of a more favorable stage distribution (i.e., more early stage cancers) at diagnosis (Table 1).

Screening adherence. Adherence is important to the overall effectiveness of a screening program. Even in a randomized trial comparing annual FOBT with no screening, only 68% of patients who were randomized to FOBT actually completed the initial screen and 63% were compliant with subsequent rescreening. Moreover, for patients with positive FOBT results, only 81% had a colonoscopy [11]. We adopted these imperfect adherence rates and assumed in the base case that adherence would be the same across strategies (Table 1).

Test performance characteristics of the CRC screening strategies. The only method for properly assessing the test performance of a given screening modality is to compare it with a reference standard in all cases. Although colonoscopy is not infallible [7], it remains the accepted gold standard for evaluating the entire colon. Therefore, the base case sensitivities and specificities for polyps and CRC for each of the screening modalities were taken from the literature following a thorough search for properly designed studies that included at least a full colonoscopy in all individuals (Table 2). For the stool-based tests and for CTC, the test performance characteristics were considered on a per person basis.

Stool-based tests. Given significant differences even between the alternative stool-based screening tests themselves (often due to different collection methods or assay types), it would not be appropriate to consider them as a class [37]. As such, we modeled different test performance scenarios for each test. A low [38] and high [14] performance level was modeled for FOBT tests that have reported in the literature (FOBT-low and FOBT-high, respectively) and a low [39,40], mid [41], and high [42] performance level was modeled for FIT assays that have been reported in the literature (FIT-low, FIT-mid, and FIT-high, respectively). The intent of modeling different levels of test performance for FOBT and FIT was to represent the range reported in the literature. This range is greatest for FIT, likely due to differences in collection methods and assays (Table 2). FIT-low represents that reported by Morikawa et al. [39,40] who studied the Magstream system with 1 d of stool collection. FIT-mid represents that reported by Nakama et al. [41] who used a 2-d method with the Monohaem system. FIT-high represents that reported by Levi et al. [42] who used the FlexSure OBT technology following 3 d of fecal collection. Both the first- [38] and second- [14] generation fecal DNA assays were modeled (FDNA-SDT1 and FDNA-SDT2, respectively).

Flexible sigmoidoscopy. Flexible sigmoidoscopy can evaluate the left colon to the splenic flexure, although this is not always possible [43]. Routine clinical practice is generally to perform a full colonoscopy in individuals found to have an

Table 1. Base case model inputs and ranges considered.

Variable	Values	Range	References
Age-dependent variables			
50- to 64-y-old individuals			
Prevalence of nonadvanced adenomas	0.171	(0.10–0.25)	[31]
Prevalence of advanced adenomas	0.038	(0.02–0.05)	[31]
Prevalence of CRC	0.001	(0.0005–0.002)	[31]
Annual death risk	0.005	—	[35]
65- to 75-y-old individuals			
Prevalence of nonadvanced adenomas	0.173	(0.10–0.25)	[31]
Prevalence of advanced adenomas	0.082	(0.05–0.10)	[31]
Prevalence of CRC	0.007	(0.002–0.01)	[31]
Annual death risk	0.018	—	[35]
Age-independent variables			
Probability of annual transition from:			
No polyp to nonadvanced adenoma – no history adenoma/CRC	0.02	(0.01–0.03)	[32][a]
No polyp to nonadvanced adenoma – history adenoma/CRC	0.038	(0.03–0.05)	[32][a]
Nonadvanced to advanced adenoma	0.019	(0.01–0.03)	[32][a]
Advanced adenoma to CRC	0.048	(0.03–0.07)	[32][a]
CRC 5-y mortality rates			
Stage I	0.068	—	[36]
Stage II	0.175	—	[36]
Stage III	0.405	—	[36]
Stage IV	0.919	—	[36]
CRC stage distributions			
In unscreened patients who develop CRC, the proportion with:			
Stage I	0.145	(0.12–0.25)	[9–11]
Stage II	0.356	(0.34–0.39)	[9–11]
Stage III	0.280	(0.23–0.32)	[9–11]
Stage IV	0.219	(0.18–0.25)	[9–11]
In patients who have CRC found using FIT, FOBT, and FDNA, the proportion with:			
Stage I	0.305	(0.29–0.33)	[9–11]
Stage II	0.318	(0.30–0.35)	[9–11]
Stage III	0.243	(0.20–0.26)	[9–11]
Stage IV	0.134	(0.10–0.15)	[9–11]
In patients who have CRC found using colonoscopy, CTC, and flex sig, the proportion with:			
Stage I	0.425	(0.41–0.50)	[38,46,68]
Stage II	0.226	(0.22–0.26)	[38,46,68]
Stage III	0.267	(0.20–0.27)	[38,46,68]
Stage IV	0.082	(0.0–0.09)	[38,46,68]
Screening adherence rates (all strategies)			
1st screen	0.68	(0.30–0.80)	[9–11]
Subsequent screens	0.63	(0.10–0.80)	[9–11]
Probability of colonoscopy after positive CTC, FOBT, FIT, FDNA, or flex sig	0.81	(0.60–0.90)	[11]
Risk of bleeding			
Colonoscopy, diagnostic	0.0003	(0.0–0.009)	[69,70]
Colonoscopy, therapeutic	0.005	(0.003–0.015)	[69–72]
Risk of perforation			
Colonoscopy, diagnostic	0.0009	(0.0005–0.002)	[70,73]
Colonoscopy, therapeutic	0.0024	(0.001–0.005)	[70,73]
Flexible sigmoidoscopy	0.0002	(0.0001–0.0004)	[32]
Risk of death after endoscopic perforation	0.049	(0.01–0.15)	[74]

Table 1. Cont.

Variable	Values	Range	References
Patient utility			
No CRC	0.91	—	[63]
Early CRC	0.74	—	[63]
Advanced CRC	0.46	—	[63]
Discount rate	0.05	—	[29]

[a]Minor adjustments were applied to the rates used in the US Multi-Society Task Force model [32] such that the total of our baseline prevalence of CRC plus the number of new CRCs developing in our natural history arm closely approximated the number of CRCs observed in the control arms of the FOBT RCTs [9–11].

doi:10.1371/journal.pmed.1000370.t001

adenomatous polyp on flexible sigmoidoscopy. As such, the sensitivity of flexible sigmoidoscopy includes the additional lesions found by colonoscopy in patients identified as having an adenoma on flexible sigmoidoscopy [44–46].

CTC and colonoscopy. Landmark studies [16,18] that employed segmental unblinding methodology [47] provided the base case test performance estimates for both CTC and colonoscopy when possible. The sensitivity and specificity of CTC for polyps ≥10 mm was taken from the National CT Colonography Trial of the American College of Radiology Imaging Network (ACRIN) [16], a large multicenter study of CTC among primarily average-risk individuals. Polyps <5 mm were not reported in this study or other large cohorts of average risk individuals. However, we optimistically assumed that the sensitivity reported for 6–9 mm polyps would be the same for all polyps <10 mm. In a sensitivity analysis we reduced the sensitivity of polyps <10 mm to that reported in a meta-analysis of CTC that included higher risk patients [48]. The sensitivity of colonoscopy for polyps ≥10 mm was taken from the study of Pickhardt et al. [18], which reported the test performance of both CTC and colonoscopy based on segmental unblinding. As this study also did not report data for polyps <5 mm, the sensitivity of colonoscopy for polyps <10 mm was taken from two back-to-back colonoscopy studies (Table 2) [49,50].

Screening-related risks. Flexible sigmoidoscopy and colonoscopy are associated with risks including bleeding, perforation, and rarely, death (Table 1). Even though CTC is less invasive than colonoscopy, colonic perforations have been reported [51–53], though many of the small CTC induced perforations diagnosed with the CT in asymptomatic individuals may not be clinically important. We assumed a low risk of CTC-induced perforation in the base case analysis [51,52], and that this would never result in death (Table 1).

Costs. ***Costs related to screening.*** All costs are reported in 2008 CAN$. The direct costs of flexible sigmoidoscopy and colonoscopy, as well as costs attributed to bleeding and perforation complications [54], were based on local estimates derived from the Calgary Health Region costing database [55] and included the nonphysician costs (capital, nursing, drugs, and cleaning costs) and the physician fees for the procedure (Table 3). CTC for primary CRC screening is not currently part of the schedule of medical benefits in any province in Canada. The direct costs of CTC were therefore conservatively assumed to be the same as that of a CT abdomen/pelvis, likely an underestimate (Table 3). We assumed that stool-based screening would be offered at a person's annual visit to their general practitioner, and as such, we only considered the cost of the screening kit and related laboratory/processing costs (Table 3).

For all screening modalities, we included the relevant patient ± caregiver time and travel costs (nonmedical costs), on the basis of available surveys for flexible sigmoidoscopy, colonoscopy, FOBT, and CTC (Table 3) [8,56,57]. The nonmedical costs of FIT and fecal DNA were assumed to be the same as FOBT. In the base

Table 2. Base case test performance characteristics for the screening modalities.

Screening Modality	Sensitivity			Specificity
	Nonadvanced Adenoma	Advanced Adenoma	Cancer	
FOBT-low [38]	0.052	0.107	0.129	0.952
FOBT-high [14]	0.030	0.074	0.500	0.980
FIT-low [39,40]	0.07	0.224	0.660	0.950
FIT-mid [41]	0.180	0.540	0.810	0.960
FIT-high [42]	0.180	0.610	0.940	0.910
Colonoscopy [18,49,50,75]	0.850	0.875	0.966	1.000
Colonoscopy after positive CTC	0.900	0.970	0.99	1.000
CTC [16]	0.760	0.900	0.966	0.890
Flexible sigmoidoscopy [44,45,46]	0.650	0.750	0.750	1.000
FDNA-SDT2 [14]	0.040	0.447	0.580	0.840
FDNA-SDT1 [38]	0.076	0.151	0.516	0.944

doi:10.1371/journal.pmed.1000370.t002

Table 3. Base case direct health care costs and nonmedical costs and ranges considered.

Variable	Values CAN$	Range CAN$	References
FOBT[a]	12	6–18	[27]
FIT	19	10–30	[76]
Colonoscopy, diagnostic[b]	857	500–1,200	[27]
Colonoscopy, therapeutic[c]	999	700–1,700	[27]
CTC	582	440–730	[27]
FDNA	336	200–500	[25]
Flex sig	650	400–900	Determined locally
Bleeding complication	3,194	(2,400–4,000)	[54]
Perforation complication	31,223	(23,500–39,000)	[54]
Total cost of managing CRC			Determined locally and [58–62]
Stage I CRC	25,049	—	
Stage II CRC	36,143	—	
Stage III CRC	96,768	—	
Stage IV CRC	134,014	—	
Nonmedical[d]			[8,56,57,77]
FOBT	36	(25–50)	
FIT	36	(25–50)	
FDNA	36	(25–50)	
Colonoscopy	308	(200–450)	
CTC	105	(100–200)	
Flex sig	105	(100–200)	

[a]FOBT: includes cost of FOBT kit (CAN$5), processing (CAN$7).
[b]Diagnostic colonoscopy: includes physician cost of diagnostic colonoscopy (CAN$327), and nonphysician cost of colonoscopy (CAN$530).
[c]Therapeutic colonoscopy: includes physician cost of therapeutic colonoscopy (CAN$401), and nonphysician cost of therapeutic colonoscopy (CAN$598).
[d]Includes patient ± caregiver time and travel costs, but excludes productivity losses [29].
doi:10.1371/journal.pmed.1000370.t003

case, we did not consider the capital costs of initiating or administering a screening program and thus assumed that screening would be opportunistic in all strategies.

Costs related to managing CRC. Existing published data on the total costs of managing patients with CRC are outdated. We assumed that the cost of surgery for CRC has remained relatively stable and thus based our surgical costs on a Canadian study reporting 1998 figures, inflation adjusted to 2008 dollars [58].

In contrast, the cost of treating CRC with chemotherapy has increased substantially because of the development of more expensive agents [22]. To estimate the cost of chemotherapy provided for advanced CRC, we used data from the Canadian Inter-Provincial Joint Oncology Drug Review (JODR) Process [59]. These estimates were the average stage-based treatment costs for chemotherapy, taking into account that not all patients would be eligible for or would comply with treatment. Patients with stage IIB disease (~50% of stage II patients) are generally managed with adjuvant chemotherapy using eight cycles of capecitabine [60,61]. First line therapy for patients with stage III CRC was assumed to be 6 mo of oxaliplatin-based therapy [62]. Considering the most recent clinical trials and assumed standards of care, the average patient with stage IV CRC received approximately 10 mo of infusional fluorouracil (5-FU), leucovorin, and oxaliplatin (FOLFOX) in combination with bevacizumab, followed by 14 doses of infusional 5-FU, leucovorin, and irinotecan (FOLFIRI). Those lacking *K-Ras* mutations were assumed to go on to receive 4 mo of anti-epidermal growth factor receptor–based inhibition therapy. We did not include the potential costs of liver metastectomy among stage IV patients, or the cost of preoperative radiation therapy in patients with operable rectal cancer.

Valuing health benefits. Health benefits were measured in terms of QALYs gained. We obtained utilities for relevant health states on the basis of a study that used a standard gamble exercise in patients with a previous history of CRC or polyps who were presented with stage-dependent outcome states for CRC (Table 1) [63].

Sensitivity Analysis

Allowance for uncertainty in the base case polyp and CRC prevalence estimates, mortality assumptions, screening test performance characteristics, screening-related risks, and costs were considered through the use of univariate and probabilistic sensitivity analyses. A number of scenario analyses were also included. We considered a scenario in which the additional costs of biologic chemotherapies for advanced stage CRC were excluded. We also examined scenarios where FIT was offered every 2 y instead of annually and analyzed our results without nonmedical costs. We assessed the impact of differential adherence rates across strategies at the initial screening encounter. For this analysis, we used the adherence rates determined by Hol et al. in a RCT comparing participation rates of FOBT, FIT, and flexible sigmoidoscopy in a screening population [13]. As Hol et al. [13] did not study fecal DNA, colonoscopy, or CTC, we assumed that fecal DNA would have the same adherence as FIT due to its comparable simplicity for patients, and we assumed that

colonoscopy would have the same adherence as flexible sigmoidoscopy. We also assessed the impact of lower subsequent adherence for the annual stool-based tests, since screening noncompliance may be more prevalent with an annual test compared to one offered less frequently. To do this assessment, we examined scenarios with decreased FOBT and FIT follow-up adherence.

Because we did not include any administrative costs for any of the CRC screening programs, we performed a sensitivity analysis to assess the impact of including administrative costs for the various screening tests. We were unable to identify a document that has reported the setup and operating costs for a population-based CRC screening program, but it is possible that programs that screen annually (i.e., stool-based tests) might have higher administrative costs than ones that screen patients every 10 y (i.e., colonoscopy). We provide sensitivity analyses varying the administrative costs per screening test between CAN$10 and CAN$50 to determine the impact on the results, making the assumption that programs screening more frequently will incur higher administrative costs.

To address limitations in classic univariate sensitivity analysis, we also performed probabilistic sensitivity analysis, which allows for the simultaneous sensitivity analysis of all variables over their plausible range [64,65]. It does so by replacing estimates of probabilities, utilities, and costs with specific probability distributions, which are based on the reported means and variances for each variable. Statistical distributions were created around all of the variables for which there was substantial measurement uncertainty, including use of a beta distribution for proportions (i.e., mortality, proportion of patients with Stage I, II, III, and IV cancer), use of a normal distribution for normally distributed variables (i.e., certain costs and utility measures), log-normal distribution for skewed variables (i.e., certain costs), and triangular distributions for variables with a range, but no statistical distribution (i.e., adenoma transition over time, probability of adherence). Given that sensitivity and specificity are linked variables that do not vary independently (linked via receiver operating curves that were unavailable), these variables were not included within the probabilistic analyses—as noted above, the sensitivity and specificity of the various screening tests were subjected to wide sensitivity analysis using the testing characteristics provided by different primary studies.

Results

Base Case Analysis

Annual CRC screening using FIT, assuming mid-range test performance characteristics, was the preferred strategy for average risk individuals in the base case analysis (Table 4). It was more effective and less costly than almost all of the other strategies including no screening. Only FIT when assuming even better test performance characteristics (i.e., FIT-high) produced more QALYs and resulted in fewer CRCs than FIT-mid, but at an additional cost of CAN$85,150 per QALY gained.

Using base case estimates, over the lifetimes of a 100,000 patient cohort, 4,857 and 1,782 individuals would develop and die from CRC, respectively, if CRC screening was not undertaken (Table 5). This "no screening" strategy would be expected to cost an average of CAN$1,901 per patient. Annual screening with FIT-mid would reduce the overall number of cancers by 71% and CRC mortality by 74% while saving CAN$68 per patient. Compared with the most effective FOBT strategy, FIT-mid would be expected to reduce the number of cancers by 60%, and CRC mortality by 63%, while saving CAN$362 per person.

Sensitivity Analysis

Under no circumstances did flexible sigmoidoscopy, FOBT, CTC, or fecal DNA appear attractive in comparison to other CRC screening modalities. As such, these strategies are not reported in our sensitivity analysis table (Table 6). Lowering the cost of CRC treatment by excluding the use of biologic chemotherapies resulted in a scenario where FIT-mid resulted in additional costs compared to no screening (CAN$163 per patient or CAN$3,691 per QALY gained). Increasing the cost of FIT testing by 50% had a similar effect; FIT-mid cost an additional CAN$105 per patient and was associated with a cost per QALY of CAN$2,375 compared to no screening. Biennial screening using FIT-mid increased the cost savings when compared to no screening. However, performing FIT less frequently also made it less effective.

When the initial adherence rates for each of the strategies was no longer assumed to be identical, FIT-mid remained dominant over no screening (Table 6). Assuming the base case initial adherence rates, when we dropped the adherence rates for subsequent screens for all of the annual fecal-based strategies, FIT-mid remained dominant over no screening. However, when subsequent adherence for FIT was dropped from 63% to 40%, both FIT-mid and FIT-high became dominant over no screening, and colonoscopy became the most effective strategy at a cost per QALY gained of CAN$300,609 compared to FIT-high. When subsequent adherence for FIT was decreased to only 20%, colonoscopy remained the most effective strategy, at a cost per QALY gained of CAN$32,912 compared to FIT-high (Table 6).

Finally, we performed a sensitivity analysis to assess the impact of higher administrative costs that might be associated with an annual screening program (i.e., FIT) compared to one offered less frequently (i.e., colonoscopy). We noted that FIT remained dominant over no screening unless the administrative costs were ~CAN$10 per test. If administrative costs were CAN$30 per test, annual FIT was associated with a cost per QALY of CAN$3,120 compared with no screening. However, if the administrative costs were CAN$50 per test, then colonoscopy would be the preferred screening modality compared with FIT, and would be associated with a cost per QALY gained of CAN$5,903 compared with no screening.

Our probabilistic sensitivity analysis revealed that FIT-mid was cost saving and more effective compared with no screening in nearly 100% of the simulations performed, confirming the robustness of the results (Figure 2).

Discussion

Our study demonstrates that annual screening with FIT, assuming mid-range test performance characteristics, is more effective and less costly than other CRC screening strategies, including the most commonly used stool-based CRC screening test, FOBT, and no screening. Among a cohort of 100,000 average risk individuals followed until death, 4,857 cancers and 1,782 cancer-related deaths would be expected with no screening. An annual FIT with high sensitivity for cancer (81%) and moderate sensitivity for advanced adenomas (54%) [41] could reduce costs and decrease the number of CRCs and cancer-related deaths to 1,393 and 457, respectively. Screening with FIT was also more effective at reducing cancer and cancer-related deaths at lower costs compared with FOBT.

FIT represents a significant advance over the traditional guaiac-based FOBTs, in large part due to FITs improved sensitivity for identifying adenomatous polyps. Our findings underscore the importance of identifying patients with advanced adenomas and

Table 4. Base case incremental cost per QALY gained for average risk patients (reported value compares strategy reported in the column with the strategy reported in the row).

Screening	Average Costs (CAN$) (95% CI)[a]	Average QALYs (95% CI)[a]	Incremental Cost Per QALY Gained										
			FIT-Mid	No Screening	FIT-High (CAN$)	FIT-Low (CAN$)	FOBT-High (CAN$)	Colonoscopy (CAN$)	FOBT-Low (CAN$)	Flex Sig (CAN$)	CTC (CAN$)	FDNA-SDT 2 (CAN$)	FDNA-SDT1 (CAN$)
FIT-mid	1,833 (1,275–1,924)	11.300 (11.29–11.30)	—	Dominated[b]	85,150	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]
No screening	1,901 (1,641–2,226)	11.255 (11.24–11.26)	—	—	2,219	3,883	15,991	4,870	18,595	10,008	12,500	25,974	82,747
FIT-high	2,004 (1,353–2,207)	11.302 (11.29–11.31)	—	—	—	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]
FIT-low	2,005 (1,519–2,020)	11.282 (11.27–11.29)	—	—	—	—	Dominated[b]	6,706	Dominated[b]	27,158	28,871	Dominated[b]	Dominated[b]
FOBT-high	2,084 (1,820–2,301)	11.267 (11.25–11.27)	—	—	—	—	—	573	25,341	7,247	11,137	36,044	Dominated[b]
Colonoscopy	2,100 (1,536–2,120)	11.296 (11.29–11.30)	—	—	—	—	—	—	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]	Dominated[b]
FOBT-low	2,195 (1,892–3,375)	11.271 (11.26–11.28)	—	—	—	—	—	—	—	3,325	8,617	42,870	Dominated[b]
Flex sig	2,263 (2,136–2,433)	11.291 (11.28–11.30)	—	—	—	—	—	—	—	—	32,489	200	Dominated[b]
CTC	2,409 (2,124–2,508)	11.296 (11.27–11.28)	—	—	—	—	—	—	—	—	—	Dominated[b]	Dominated[b]
FDNA-SDT2	2,491 (2,187–2,644)	11.278 (11.27–11.28)	—	—	—	—	—	—	—	—	—	—	Dominated[b]
FDNA-SDT1	2,720 (2,422–2,937)	11.265 (11.25–11.27)	—	—	—	—	—	—	—	—	—	—	—

[a]95% confidence intervals (CIs) based on probabilistic sensitivity analysis using baseline statistical distributions around all uncertain variables.
[b]Dominated is defined as more costly and fewer QALYs compared with the strategy reported in the row.
doi:10.1371/journal.pmed.1000370.t004

Table 5. Cancer outcomes and number of screening tests required during the lifetimes for a hypothetical 100,000 average risk patient cohort.

Screening Test	*n* Cancers Overall[a]	*n* Cancer Deaths	*n* Primary Screening Tests	*n* Colonoscopies	Cost Of Screening And Managing CRC (CAN$)
FIT-high	1,290	432	819,178	56,541	2,004
FIT-mid	1,393	457	822,077	53,909	1,833
CTC	1,796	593	188,315	58,354	2,409
Colonoscopy	1,825	624	155,210	N/A	2,100
Flex Sig	2,036	699	189,135	49,484	2,263
FIT-low	2,634	918	871,986	31,597	2,005
FDNA-SDT2	3,129	1,143	331,090	20,805	2,491
FOBT-low	3,457	1,250	889,168	21,805	2,195
FOBT-high	3,890	1,368	902,299	15,089	2,084
FDNA-SDT1	4,131	1,530	331,699	14,548	2,720
No screening	4,857	1,782	n/a	n/a	1,901

[a]*n* cancers overall include symptomatic and screen found CRC.

doi:10.1371/journal.pmed.1000370.t005

preventing cancer through the identification and removal of precancerous polyps. Indeed, changing the sensitivity of FIT for cancer had relatively little impact on our results, whereas reducing the sensitivity of FIT for advanced adenomas from 54% to below 45% resulted in FIT no longer being cost saving compared with no screening.

Although it may seem counter-intuitive that screening with FIT could be even more effective than colonoscopy, this is due to the more frequent screening interval with FIT. In base case analyses, and consistent with current guidelines [3,6], screening with FIT was done annually compared to every 10 y with colonoscopy. Therefore, even though the test performance of a single FIT test was inferior to colonoscopy, there were more opportunities to identify previously missed pathology with FIT compared to colonoscopy.

Our results are robust. FIT with mid-range performance (FIT-mid) remained optimal compared with no screening and all the other strategies except FIT with even better test performance (FIT-high) unless the cost of CRC treatment was reduced, or the sensitivity for advanced adenomas was decreased significantly. However, even with lower CRC treatment costs, FIT remained economically attractive. Many health jurisdictions now fund biologic chemotherapies for advanced-stage CRC and with further advances in CRC chemotherapy, it is unlikely that management costs for CRC will decrease [22]. In addition, our modeled CRC treatment costs were lower than those used in a recent US study that had similar results [25], lending further support to the notion that CRC screening can indeed save money.

It is possible that the administrative costs of annual screening programs such as FIT would be more expensive over the long-run compared with those offered every 5 or 10 y. As these data are not known, we did not consider administrative costs or the costs to build and staff additional screening centers in our primary analysis. However, in sensitivity analysis, we noted that FIT-mid remained cost saving if the administration costs were <CAN$10 per test, and remained attractive compared with colonoscopy even if the administrative costs per test were CAN$30 per test. It should also be noted that the additional infrastructure required to implement primary screening with CTC, flexible sigmoidoscopy, or colonoscopy would likely counterbalance a substantial portion of these additional administrative costs of an annual screening program.

We assumed in the base case that adherence would be identical across all of the CRC screening strategies. Although this may not be true, we are unaware of a study that has evaluated screening uptake for all of the strategies we considered. However, fecal-based screening does not require a bowel preparation, is associated with lower patient-borne costs, and is safe to perform, which may be more appealing to the general population. Furthermore, FIT does not require any dietary restrictions. Indeed, in a recent randomized trial, FIT was associated with higher screening uptake than flexible sigmoidoscopy and FOBT [13]. Of course, this finding only strengthens our conclusions as illustrated in our scenario analysis in which FIT had relatively higher adherence than all of the other strategies (Table 6). Recent data suggest that screening adherence with FOBT may drop by 50% after only 2 y in a biennial screening program [66]; this may affect programs with frequent screening (i.e., annual fecal-based strategies) to a greater extent than programs requiring less frequent screening (i.e., colonoscopy). As expected, when we dropped our subsequent adherence rates for FOBT and FIT, FIT-mid became less effective, though it remained dominant compared with no screening. In contrast, colonoscopy became the most effective strategy when the subsequent adherence rates for FOBT and FIT were dropped from 63% to 40%, though it was associated with an unattractive incremental cost per QALY. It is clear that further information on long-term adherence rates for annual stool-based tests are needed.

Our study has limitations. As with most economic evaluations, our results are limited by available evidence. The natural history of adenomas and their progression to cancer is not clearly known. However, we populated our model with the best available evidence including a systematic review of adenoma and CRC prevalence rates [31] and modeled new adenoma growth and adenoma progression over time to closely match high quality clinical datasets [9–11]. We did not model cancers arising from lesions other than adenomas. However, most CRCs arising in average risk individuals are believed to develop via the traditional adenoma-carcinoma sequence. A small proportion of CRC may develop from undetectable lesions (i.e., flat or depressed adenomas), and it is known that some interval cancers can arise through a rapid adenoma-carcinoma sequence between screening studies [67]. It should be noted that this potential issue would

Table 6. Sensitivity analysis.

Screening	Cost of Screening and Management (CAN$)[a]	QALY	Incremental Cost per QALY Gained (CAN$)[a,b]
Base case			
FIT-mid	1,833	11.300	
No screening	1,901	11.255	(Dominated)[c]
FIT-high	2,004	11.302	84,876
Colonoscopy	2,100	11.296	(Dominated)[c]
Lower stage III and IV cancer costs, including chemotherapy, but without biologics (Stage II CAN$35844, Stage III CAN$80,345, and stage IV CAN$99,574)			
No screening	1,582	11.255	
FIT-mid	1,745	11.300	3,691
FIT-high	1,842	11.302	89,921
Colonoscopy	1,990	11.296	(Dominated)[c]
Increase FIT direct cost by 50%			
No screening	1,901	11.255	
FIT-mid	2,006	11.300	2,375
Colonoscopy	2,100	11.296	(Dominated)[c]
FIT-high	2,177	11.302	84,750
Biennial FIT screening (versus annual FIT screening modeled in baseline analyses)			
FIT-mid	1,736	11.289	
FIT-high	1,784	11.291	19,606
No screening	1,901	11.255	(Dominated)[c]
Colonoscopy	2,100	11.296	64,741
Initial adherence 60% for FIT and fecal DNA, 50% for FOBT, 40% for CT colonoscopy, and 30% for colonoscopy [13]			
FIT-mid	1,815	11.299	
No screening	1,901	11.255	(Dominated)[c]
FIT-high	1,986	11.301	85,927
Colonoscopy	2,055	11.279	(Dominated)[c]
Decrease subsequent adherence rates for FITs and FOBTs from 63% to 40%			
FIT-mid	1,751	11.293	
FIT-high	1,839	11.295	38,536
No screening	1,901	11.255	(Dominated)[c]
Colonoscopy	2,100	11.296	300,609
Decrease subsequent adherence rates for FITs and FOBTs from 63% to 20%			
FIT-mid	1,752	11.283	
FIT-high	1,772	11.286	8,709
No screening	1,901	11.255	(Dominated)[c]
Colonoscopy	2,100	11.296	32,912
CAN$10 administrative cost added for all screening tests			
No screening	1,901	11.255	
FIT-mid	1,902	11.300	17
FIT-high	2,075	11.302	85,831
Colonoscopy	2,109	11.296	(Dominated)[c]
CAN$50 administrative cost added for all screening tests			
No screening	1,901	11.255	
Colonoscopy	2,143	11.296	5,903
FIT-mid	2,176	11.300	10,202
FIT-high	2,357	11.302	89,651

[a]Numbers rounded to nearest CAN$1.
[b]Each incremental value compares the value of that strategy to next most costly, nondominated, strategy.
[c]Dominated is defined as more costly and fewer QALYs compared with a comparator strategy.
doi:10.1371/journal.pmed.1000370.t006

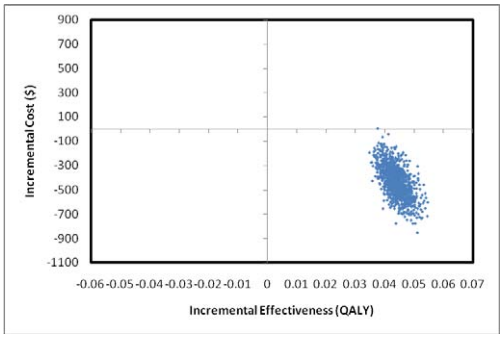


Figure 2. Probabilistic sensitivity analysis. An incremental cost-effectiveness scatterplot comparing FIT-mid with no screening in which the uncertainty in all model inputs has been tested simultaneously. Data points in the lower right quadrant reflect situations where FIT-mid is more effective and less costly than no screening.
doi:10.1371/journal.pmed.1000370.g002

impact the effectiveness of all CRC screening modalities, and thus would be unlikely to impact the differential effectiveness between our modeled strategies. Given data limitations, we modeled identical CRC stage distributions for cancers detected using all of the stool-based strategies despite differences in testing characteristics. Given FIT's superior sensitivity compared to FOBT, patients diagnosed with CRC might be expected to have more earlier stage cancers, which again would make FIT appear more attractive. We assumed that the results of each screening test were independent of the prior test result. While not informed by evidence, it is possible that this is not entirely true; however, it is important to note that the results of our analysis were robust to small changes in the sensitivity and specificity of each of the screening tests. Finally, although we did model the most widely available and promising screening strategies, additional technologies are being developed and it is possible that other screening paradigms, including nurse-based endoscopy, may become viable in the future as a means to reduce the cost of delivering flexible sigmoidoscopy and potentially colonoscopy.

In conclusion, annual screening with FIT having test performance characteristics within the mid-range reported in the literature is both more effective and less costly than other CRC screening modalities, including FOBT and colonoscopy, and not screening for CRC. Even if this level of test performance is not attainable in clinical practice, annual screening with a lower performing FIT is still highly attractive with a cost per QALY gained of <CAN$5,000 compared to no screening. Our results are robust suggesting that screening for CRC with FIT should be considered the modality of choice for average risk patients between the ages of 50 and 75 in North America.

Author Contributions

ICMJE criteria for authorship read and met: SJH RJH FA SD BJM. Agree with the manuscript's results and conclusions: SJH RJH FA SD BJM. Designed the experiments/the study: SJH RJH BJM. Analyzed the data: SJH FA BJM. Collected data/did experiments for the study: SJH FA SD BJM. Wrote the first draft of the paper: SJH FA. Contributed to the writing of the paper: RJH FA SD BJM.

References

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, et al. (2007) Cancer statistics, 2007. CA Cancer J Clin 57: 43–66.
2. Wilson JMG, Junger G (1968) Principles and practice of screening for disease. Geneva: WHO.
3. (2008) Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 149: 627–637.
4. Leddin D, Hunt R, Champion M, Cockeram A, Flook N, et al. (2004) Canadian Association of Gastroenterology and the Canadian Digestive Health Foundation: Guidelines on colon cancer screening. Can J Gastroenterol 18: 93–99.
5. Winawer S, Fletcher R, Rex D, Bond J, Burt R, et al. (2003) Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence. Gastroenterology 124: 544–560.
6. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, et al. (2008) 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 134: 1570–1595.
7. Lieberman D (2004) Colonoscopy: as good as gold? Ann Intern Med 141: 401–403.
8. Heitman SJ, Au F, Manns BJ, McGregor SE, Hilsden RJ (2008) Nonmedical costs of colorectal cancer screening with the fecal occult blood test and colonoscopy. Clin Gastroenterol Hepatol 6: 912-917 e911.
9. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, et al. (1996) Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 348: 1472–1477.
10. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O (1996) Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 348: 1467–1471.
11. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, et al. (1993) Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med 328: 1365–1371.
12. van Rossum LG, van Rijn AF, Laheij RJ, van Oijen MG, Fockens P, et al. (2008) Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 135: 82–90.
13. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, et al. (2010) Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 59: 62–68.
14. Ahlquist DA, Sargent DJ, Loprinzi CL, Levin TR, Rex DK, et al. (2008) Stool DNA and occult blood testing for screen detection of colorectal neoplasia. Ann Intern Med 149: 441-450, W481.
15. Itzkowitz SH, Jandorf L, Brand R, Rabeneck L, Schroy PC, 3rd, et al. (2007) Improved fecal DNA test for colorectal cancer screening. Clin Gastroenterol Hepatol 5: 111–117.
16. Johnson CD, Chen MH, Toledano AY, Heiken JP, Dachman A, et al. (2008) Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 359: 1207–1217.
17. Kim DH, Pickhardt PJ, Taylor AJ, Leung WK, Winter TC, et al. (2007) CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 357: 1403–1412.
18. Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, et al. (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 349: 2191–2200.
19. Heitman SJ, Manns BJ, Hilsden RJ, Fong A, Dean S, et al. (2005) Cost-effectiveness of computerized tomographic colonography versus colonoscopy for colorectal cancer screening. Cmaj 173: 877–881.
20. Pickhardt PJ, Hassan C, Laghi A, Zullo A, Kim DH, et al. (2007) Cost-effectiveness of colorectal cancer screening with computed tomography colonography: the impact of not reporting diminutive lesions. Cancer 109: 2213–2221.
21. Vijan S, Hwang I, Inadomi J, Wong RK, Choi JR, et al. (2007) The cost-effectiveness of CT colonography in screening for colorectal neoplasia. Am J Gastroenterol 102: 380–390.
22. Schrag D (2004) The price tag on progress--chemotherapy for colorectal cancer. N Engl J Med 351: 317–319.
23. Benson VS, Patnick J, Davies AK, Nadel MR, Smith RA, et al. (2008) Colorectal cancer screening: a comparison of 35 initiatives in 17 countries. Int J Cancer 122: 1357–1367.
24. Pignone M, Saha S, Hoerger T, Mandelblatt J (2002) Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 137: 96–104.
25. Parekh M, Fendrick AM, Ladabaum U (2008) As tests evolve and costs of cancer care rise: reappraising stool-based screening for colorectal neoplasia. Aliment Pharmacol Ther 27: 697–712.
26. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, et al. (2008) Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med 149: 659–669.
27. Ho C, Heitman S, Membe SK, Morrison A, Moulton K, et al. (2008) Computed tomographic colonography for colorectal cancer screening in an average risk population: systematic review and economic evaluation. Canadian Agency for Drugs and Technologies in Health, ed. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health.
28. Rabeneck L, Paszat LF (2004) A population-based estimate of the extent of colorectal cancer screening in Ontario. Am J Gastroenterol 99: 1141–1144.

29. (2006) Guidelines for the economic evaluation of health technologies. Canadian Agency for Drugs and Technologies in Health, ed. 3rd edition Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health.
30. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, et al. (2003) Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. Value Health 6: 9–17.
31. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, et al. (2009) Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 7: 1272–1278.
32. Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, et al. (1997) Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 112: 594–642.
33. Pickhardt PJ, Choi JR, Hwang I, Schindler WR (2004) Nonadenomatous polyps at CT colonography: prevalence, size distribution, and detection rates. Radiology 232: 784–790.
34. Lewis JD, Ng K, Hung KE, Bilker WB, Berlin JA, et al. (2003) Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy. Arch Intern Med 163: 413–420.
35. (2007) Death, by age group and sex, Canada, provinces and territories, annual. Ottawa, Canada: Statistics Canada.
36. O'Connell JB, Maggard MA, Ko CY (2004) Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst 96: 1420–1425.
37. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R (2008) Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 149: 638–658.
38. Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA, Ross ME (2004) Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 351: 2704–2714.
39. Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, et al. (2007) Sensitivity of immunochemical fecal occult blood test to small colorectal adenomas. Am J Gastroenterol 102: 2259–2264.
40. Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, et al. (2005) A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. Gastroenterology 129: 422–428.
41. Nakama H, Zhang B, Zhang X (2001) Evaluation of the optimum cut-off point in immunochemical occult blood testing in screening for colorectal cancer. Eur J Cancer 37: 398–401.
42. Levi Z, Rozen P, Hazazi R, Vilkin A, Waked A, et al. (2007) A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 146: 244–255.
43. Painter J, Saunders DB, Bell GD, Williams CB, Pitt R, et al. (1999) Depth of insertion at flexible sigmoidoscopy: implications for colorectal cancer screening and instrument design. Endoscopy 31: 227–231.
44. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, et al. (2000) Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. N Engl J Med 343: 169–174.
45. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, et al. (2003) Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. Ann Intern Med 139: 959–965.
46. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, et al. (2000) Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. N Engl J Med 343: 162–168.
47. Pineau BC, Paskett ED, Chen GJ, Durkalski VL, Espeland MA, et al. (2001) Validation of virtual colonoscopy in the detection of colorectal polyps and masses: rationale for proper study design. Int J Gastrointest Cancer 30: 133–140.
48. Mulhall BP, Veerappan GR, Jackson JL (2005) Meta-analysis: computed tomographic colonography. Ann Intern Med 142: 635–650.
49. Hixson LJ, Fennerty MB, Sampliner RE, McGee D, Garewal H (1990) Prospective study of the frequency and size distribution of polyps missed by colonoscopy. J Natl Cancer Inst 82: 1769–1772.
50. Rex DK, Cutler CS, Lemmel GT, Rahmani EY, Clark DW, et al. (1997) Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. Gastroenterology 112: 24–28.
51. Burling D, Halligan S, Slater A, Noakes MJ, Taylor SA (2006) Potentially serious adverse events at CT colonography in symptomatic patients: national survey of the United Kingdom. Radiology 239: 464–471.
52. Pickhardt PJ (2006) Incidence of colonic perforation at CT colonography: review of existing data and implications for screening of asymptomatic adults. Radiology 239: 313–316.
53. Sosna J, Blachar A, Amitai M, Barmeir E, Peled N, et al. (2006) Colonic perforation at CT colonography: assessment of risk in a multicenter large cohort. Radiology 239: 457–463.
54. Heitman SJ, Fong A, Rabeneck L, Manns BJ, Hilsden RJ (2005) Direct costs of bleeding and perforation complications from colonoscopy. Am J Gastroenterol 100: S281.
55. Lee H, Manns B, Taub K, Ghali WA, Dean S, et al. (2002) Cost analysis of ongoing care of patients with end-stage renal disease: the impact of dialysis modality and dialysis access. Am J Kidney Dis 40: 611–622.
56. Frew E, Wolstenholme JL, Atkin W, Whynes DK (1999) Estimating time and travel costs incurred in clinic based screening: flexible sigmoidoscopy screening for colorectal cancer. J Med Screen 6: 119–123.
57. Hilsden R, Heitman S, Au F, Manns B, McGregor E (2008) Non-medical costs of colorectal cancer screening using computed tomographic colonography. Am J Gastroenterol 103.
58. Maroun J, Ng E, Berthelot JM, Le Petit C, Dahrouge S, et al. (2003) Lifetime costs of colon and rectal cancer management in Canada. Chronic Dis Can 24: 91–101.
59. (2008) Guidelines: Committee to Evaluate Drugs - Cancer Care Ontario Joint Drug Review Process, Ontario Drug Benefit Plan. Toronto, Canada: Ontario Ministry of Health.
60. Kerr DJ, Gray R, McConkey C, Barnwell J (2000) Adjuvant chemotherapy with 5-fluorouracil, L-folinic acid and levamisole for patients with colorectal cancer: non-randomised comparison of weekly versus four-weekly schedules--less pain, same gain. QUASAR Colorectal Cancer Study Group. Ann Oncol 11: 947–955.
61. Scheithauer W, McKendrick J, Begbie S, Borner M, Burns WI, et al. (2003) Oral capecitabine as an alternative to i.v. 5-fluorouracil-based adjuvant therapy for colon cancer: safety results of a randomized, phase III trial. Ann Oncol 14: 1735–1743.
62. Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, et al. (2004) Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med 350: 2343–2351.
63. Ness RM, Holmes AM, Klein R, Dittus R (1999) Utility valuations for outcome states of colorectal cancer. Am J Gastroenterol 94: 1650–1657.
64. Briggs A, Fenn P (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ 7: 723–740.
65. Briggs A, Sculpher M, Buxton M (1994) Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health Econ 3: 95–104.
66. Fenton JJ, Elmore JG, Buist DS, Reid RJ, Tancredi DJ, et al. (2010) Longitudinal adherence with fecal occult blood test screening in community practice. Ann Fam Med 8: 397–401.
67. Sawhney MS, Farrar WD, Gudiseva S, Nelson DB, Lederle FA, et al. (2006) Microsatellite instability in interval colon cancers. Gastroenterology 131: 1700–1705.
68. Regula J, Rupinski M, Kraszewska E, Polkowski M, Pachlewski J, et al. (2006) Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia. N Engl J Med 355: 1863–1872.
69. Kavic SM, Basson MD (2001) Complications of endoscopy. Am J Surg 181: 319–332.
70. Warren JL, Klabunde CN, Mariotto AB, Meekins A, Topor M, et al. (2009) Adverse events after outpatient colonoscopy in the Medicare population. Ann Intern Med 150: 849–857, W152.
71. Rosen L, Bub DS, Reed JF, 3rd, Nastasee SA (1993) Hemorrhage following colonoscopic polypectomy. Dis Colon Rectum 36: 1126–1131.
72. Waye JD, Lewis BS, Yessayan S (1992) Colonoscopy: a prospective report of complications. J Clin Gastroenterol 15: 347–351.
73. Misra T, Lalor E, Fedorak RN (2004) Endoscopic perforation rates at a Canadian university teaching hospital. Can J Gastroenterol 18: 221–226.
74. Garbay JR, Suc B, Rotman N, Fourtanier G, Escat J (1996) Multicentre study of surgical complications of colonoscopy. Br J Surg 83: 42–44.
75. Bressler B, Paszat LF, Chen Z, Rothwell DM, Vinden C, et al. (2007) Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis. Gastroenterology 132: 96–102.
76. (2009) Primary screening programs for colorectal cancer in Canada using non-invasive home tests. http://www.cadth.ca/media/pdf/hta_htupdate_issue-9_e.pdf. Accessed 2 January, 2009.
77. Jonas DE, Russell LB, Sandler RS, Chou J, Pignone M (2008) Value of patient time invested in the colonoscopy screening process: time requirements for colonoscopy study. Med Decis Making 28: 56–65.

Editors' Summary

Background. Colorectal (bowel) cancer is the second leading cause of cancer deaths for both men and women in North America. Colorectal cancer screening is an important means for reducing morbidity and mortality and fulfils the World Health Organization criteria for mass screening. However, a variety of CRC screening approaches are available. Colonoscopy is viewed as the gold standard of colorectal cancer screening as it has a high sensitivity for identifying adenomas and cancer and polyps can be removed during the screening examination. However, colonoscopy is associated with a number of complications and there are also barriers to access. Another type of test, the guaiac fecal occult blood test, has been shown to reduce mortality from colorectal cancer but this test has low sensitivity for identifying colorectal neoplasia, particularly adenomas. Fecal immunochemical tests, which also detect blood in the stool, have improved test performance characteristics (high sensitivity and specificity) and the potential to improve participation rates compared to guaiac fecal occult blood test and flexible sigmoidoscopy. Fecal DNA (a stool test, based on the detection of DNA shed by cancerous tissue) is another screening option, as is computed tomographic colonography ("virtual" colonoscopy), that might rival colonoscopy in detecting advanced adenomas and colorectal cancer but is expensive and requires a full colonic preparation.

Why Was This Study Done? In the absence of firm comparative evidence to guide the selection of any one screening modality and given the varied test performance characteristics and the significant differences in costs and resources associated with each, a robust cost-effectiveness analysis might help health policy makers in deciding whether or not to offer screening and if so, in selecting the most appropriate and cost effective screening modality. In this study the researchers conducted a full economic evaluation of all relevant colorectal cancer screening modalities in North America.

What Did the Researchers Do and Find? The researchers used an incremental cost-utility analysis, a sophisticated modeling technique, and two hypothetical patient cohorts (individuals with an "average risk," i.e., no family history of colorectal cancer, aged 50–64 and 65–75) to compare guaiac-based fecal occult blood test or fecal immunochemical test annually (the researchers considered three distinct fecal immunochemical testing strategies on the basis of assays and collection methods taken from studies that have reported "low," "mid," and "high" test performance characteristics), fecal DNA every three years, flexible sigmoidoscopy or computed tomographic colonography every 5 years, and colonoscopy every 10 years. The researchers also included a no screening natural history arm as a comparison to each screening approach. For the baseline data of their model, the researchers used adenoma and colorectal prevalence rates from a recent systematic review and based screening adherence, test performance, and colorectal treatment costs on available data. The researchers found that annual fecal immunochemical testing with mid-range testing characteristics, was more effective and less costly compared to all strategies (including no screening). Using this screening modality, among the lifetimes of 100,000 average-risk patients, the number of cancers could be reduced from 4,857 to 1,393 and the number of deaths from colorectal cancer from 1,782 to 457, while saving CAN$68 per person. Although in the sensitivity and scenario analysis, screening patients using fecal immunochemical testing became more expensive than a strategy of no screening when the test performance of fecal immunochemical testing was reduced, or the cost of managing colorectal cancers was lowered, the researchers found that screening for colorectal cancer with fecal immunochemical testing remained the most economically attractive screening option.

What Do These Findings Mean? This model-based economic analysis found that fecal immunochemical testing is more effective and less costly than all other colorectal screening strategies, including the most commonly-used stool-based screening test, guaiac-based fecal occult blood testing, and no screening. Furthermore, this study suggests that annual screening with fecal immunochemical testing (assuming mid-range test performance characteristics) reduces the risk of colorectal cancer and colorectal cancer–related deaths, and lowers health care costs in comparison to all other screening strategies and to no screening. Therefore, health policy makers should consider prioritizing funding for fecal immunochemical testing as the screening modality for colorectal cancer.

Additional Information. Please access these Web sites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed.1000370.

- Cancer.org has information for patients on colorectal cancer
- The US Centers for Disease Control (CDC) list colorectal screening guidelines
- The CDC also provides patient information on colorectal cancer Screening

CHAPTER 4: DISCUSSION

4.0 Summary of Research Findings

The findings generated from this thesis not only address several knowledge gaps and policy-relevant questions, but also provide support for investing in population-based CRC screening using FIT for average risk people aged 50-75. Furthermore, the model we have developed will be an important resource for conducting future economic evaluations of alternative screening strategies (e.g. delaying screening beyond 50, nurse led colonoscopy) and of new CRC screening technologies as they become available.

In our meta-analysis of North American observational studies [2] (Chapter 2) we showed that the prevalence of CRC among average risk individuals is 0.3%. Consistent with the known relationship between age and risk of CRC, we found that the prevalence of CRC among those under age 65 was 0.1%, rising to 0.7% among those over age 65. Important on its own, however was the finding that the prevalence of advanced adenomas was 5.7% among average risk individuals. The prevalence of advanced adenomas similarly increased with age (3.8% for those < 65; 8.2% for those ≥ 65). These values were used as estimates of the prevalence of CRC and advanced adenomas for a 50-64 year old cohort (mean age 58) and a 65-75 year old cohort (mean age 70), respectively and suggest that 1 out of every 150 and 1 out of every 12 individuals age 65-75 harbors prevalent CRC or a lesion likely to develop into a cancer, respectively. Nevertheless, despite effective screening options, the majority of these CRC prevalent cases and at-risk individuals will remain undetected given that the majority of eligible individuals are not currently being screened.

Although the cost of mass screening for CRC has the potential to be high, the competing cost of treating advanced-stage CRC has been rapidly rising over the past decade [25]. In our economic evaluation of CRC screening [3] (Chapter 3) we showed using a decision analytic model that screening for CRC with FIT not only saves lives, but also offers the potential to *save* money. Annual FIT, assuming mid-range testing characteristics, was more effective and less costly compared to all strategies (including no screening) except FIT offering even superior test performance. Among the lifetimes of 100,000 average-risk individuals, the number of cancers could be reduced from 4,857 to 1,393 and the number of CRC deaths from 1,782 to 457, while saving $68 Canadian dollars per person. Although screening with FIT would require over 800,000 tests per 100,000 people screened over a lifetime, our model also suggests that the number of necessary colonoscopies could be lowered by over 100,000. Specifically, primary screening with colonoscopy would require 155,000 colonoscopies compared to 54,000 colonoscopies in a FIT-based program among a 100,000 patient cohort of average risk individuals.

While the findings generated from our model were generally robust to sensitivity analyses, we noted that FIT was no longer dominant over the other screening strategies when its test performance was substantially reduced (to that of FIT low) or when its cost was increased by 50%, and when the cost of managing CRC was lowered dramatically through eliminating the cost of biological chemotherapies. Nevertheless, even under these somewhat unlikely circumstances, FIT remained economically attractive with a cost/QALY < $5000. Finally, when the overall model uncertainly was assessed in our

probabilistic sensitivity analysis, FIT-mid remained dominant over no screening in nearly 100% of the simulations performed.

4.1 Considerations for Health Care Decision Makers and Public Policy

Effectiveness and cost-effectiveness are only two of the elements considered by those who determine how to allocate finite health care budgets. Other important factors include the size of the budget and the prevailing health care priorities. For instance, even if screening for CRC is cost-effective, it wouldn't be an option if funding it would overwhelm the budget. Furthermore, regardless of how economically attractive an intervention is, it might be overlooked if the health care issue in question wasn't considered to be important. However, CRC is a common deadly cancer that killed nearly 9,000 Canadians in 2011 [5]. As well, our results indicate that the health care costs associated with not screening for CRC at least rival the costs of screening using FIT, and that we may in fact be wasting a portion of our scarce health care dollars by not funding FIT for average risk individuals.

It could be argued that some of our model assumptions (i.e. disease prevalence, screening test performance, screening adherence, costs) may have been overly optimistic to the extent that screening with FIT may not be cost saving in reality. Nevertheless, as our extensive sensitivity analyses demonstrate, screening average risk individuals with FIT is at a minimum highly attractive from an economic perspective compared to other currently funded health care interventions. On the other hand, if saving money were the only goal, another option might be to scale back funding for treatment of CRC and specifically to limit or even eliminate the use of expensive biologic chemotherapies.

Although a consideration, it seems unlikely that society would support such a policy, since several new expensive drugs for CRC have recently been funded and are currently in use (e.g. bevacizumab, panitumumab). Instead, it seems more likely that our health care system will face rising costs for chemotherapeutic agents. For instance, Regorafenib, a new biologic chemotherapy drug for patients with metastatic colorectal carcinoma, was shown to increase survival by 1.4 months compared to standard care among those with stage-IV disease [53]. One can reasonably expect that the incremental cost of Regorafenib will be high, and if past decisions are followed, it is likely that this drug will eventually be reimbursed in Canada.

It is encouraging that several countries including the United Kingdom and a few in Europe are in the process of implementing national population CRC screening programs. At present there are only a few province-wide CRC screening programs in Canada. “Colon Cancer Check” was launched by Cancer Care Ontario in 2008. This program targets average risk individuals age 50-74 with FOBT every 2 years. According to Cancer Care Ontario’s website, the target screening uptake was to be 40% by 2011. They cite that screening adherence has increased from 15% in 2003-04 to 30% in 2008-09 [54]. Unfortunately, there is no organized provincial CRC screening program in Alberta. Plans for an Alberta Colorectal Cancer Screening Program (ACRCSP) offering FIT every 2 years for average risk 50-74 year olds exists, but a firm funding commitment has not been announced. Other provinces including Manitoba, British Columbia and Nova Scotia have screening programs in various phases of implementation [52].

CRC screening in Alberta continues to be opportunistic, and is generally coordinated by primary care physicians [55]. As such, only those patients who are informed

and self-motivated or under the care of a physician who considers CRC a priority are likely to participate in screening. Thus, the capacity to reduce disease-specific mortality through prevention (i.e. polypectomy) and early detection of subclinical CRC is by definition, limited. Indeed, the uptake of opportunistic screening in Alberta has been very low [56]. A number of barriers to potentially account for this have been identified including low patient awareness [57], limited access to endoscopy [55] and high patient-borne costs [58]. There are a few ad hoc CRC screening projects in Calgary (Forzani and MacPhail Colon Cancer Screening Centre – CCSC), Edmonton (Stop Colorectal Cancer through Prevention and Education – SCOPE Project) and Lethbridge (Lethbridge and Area Colorectal Cancer Screening Program). Despite the growing problem in supplying timely access to endoscopy services across Canada and particularly in Alberta [59], these initiatives generally focus on colonoscopy for CRC screening.

A province-wide CRC screening program that offers a single safe, effective and cost-effective entry strategy (i.e. FIT) has a number of other potential advantages over and above those supported by our model results. A provincially administered FIT-based program could reduce patient and provider confusion regarding appropriate screening. By matching the FIT test-positive threshold and program adherence targets to the availability of colonoscopy resources (arguably the most scarce resource), such a program could succeed in delivering the needed follow-up (i.e. colonoscopy) in a timely manner. In addition, a province-wide centrally administered FIT-based program could ensure that screening is occurring appropriately throughout the Province and would be well positioned to monitor the safety and quality of colonoscopy through a formal quality assurance process. Finally, it is likely that a province-wide screening program would

offer the best chance of improving CRC screening uptake, which is of central importance to lowering the overall burden of this disease.

The literature has continued to evolve since the publication of our manuscripts. A recent randomized trial in a population-based CRC screening setting compared FIT at screening at intervals of 1, 2 and 3 years. Each patient was invited for two rounds of screening. The detection of advanced neoplasia was lower in the second round and did not depend on the interval length [60]. Furthermore, participation rates were over 60% in all 3 arms and were not impacted by the screening interval. The authors suggested that screening intervals could be tailored to local resources. Although we found in a sensitivity analysis (Table 6 of the PLoS Medicine paper) that biennial FIT screening was slightly less effective than annual screening with FIT, it was also less costly. However, biennial FIT screening was still dominant over no screening and colonoscopy, although more effective than a biennial FIT strategy, was associated with a cost/QALY of $64,741. Perhaps the most significant development in the FIT literature has been the recent interim report of the landmark RCT of colonoscopy versus biennial FIT for CRC screening [61]. Although this multicenter trial involving over 50,000 patients will not be complete until 2021, the results of the first round of screening found no difference in the rate of detection of CRC between the colonoscopy and FIT screened groups and the participation rate of FIT was significantly higher than colonoscopy. More adenomas were found with colonoscopy, but these results are based on a single round of screening. Thus, it is highly probable that additional cases of adenomas and CRC will be found with repeat screening over the next decade.

Screening for CRC will never be 100% effective. Testing is imperfect and many patients will remain non-compliant even within the context of a population-based screening program. Furthermore, some patients will continue to develop cancers outside of the recommend age ranges for screening and surveillance. Nevertheless, screening appears effective in reducing the incidence of CRC and in lowering CRC-related mortality and our research suggests that screening with FIT may have little impact on overall costs in light of the current climate of expensive treatments for CRC. Thus, when our findings and the recent clinical trials of FIT are considered collectively, it seems clear that health care decision makers need to prioritize population-based CRC screening and fund FIT for average risk individuals.

4.2 Limitations

The published findings contained in this thesis represent important contributions to the literature. However, there are limitations to our studies and knowledge gaps remain that warrant further research.

The estimates of CRC and adenoma prevalence reported in our systematic review are based on the currently available North American literature. Despite the enhanced power of meta-analysis, we were only able to stratify by a relatively wide age range. Furthermore, we had insufficient power to stratify according to gender or other potential factors known to influence the risk of CRC such as BMI and smoking status. Using a cut-off of 65, we were able to identify significant differences in lesion prevalence for a 50-64 year old cohort (mean age 58) and a 65-75 year old cohort (mean age 70). However, it seems plausible that the difference in CRC and adenoma prevalence between a 75 year

old and a 50 year old could be even larger. Hence, the preferred screening strategy for a 75 year old faced with a final screening examination might be different than for a 50 year old entering a period of multiple years of follow-up. Nevertheless, as it currently stands FIT was found to be the preferred primary CRC screening strategy for both the "average" 50-64 year old and 65-75 year old individual. However, we must be open to the fact that this "one size fits all" recommendation may not hold up as more data become available on the prevalence of colorectal neoplasia across the spectrum of age and among subgroups of average risk individuals.

The Markov model we created is intended to simulate "real life" based on what is currently known on the prevalence and natural history of colorectal neoplasia, screening adherence rates, and the stage-based treatment outcomes and cost of managing CRC. The structure of the model was heavily influenced by the current CRC screening guidelines that are intended to guide clinical practice. Considerable effort was made to ensure the internal validity of our model by comparing our model outputs with published clinical datasets. This is one of the strengths of this thesis work as many economic evaluations including those performed in the context of colorectal cancer screening are based on models that haven't been formally validated. A more detailed description of the steps used to validate our model can be found in Appendix 4. Finally, we populated the model with the best available evidence including a systematic review and meta-analysis and conducted sensitivity analyses on the key model inputs driving our results and the inputs where precise estimates were unavailable. Therefore, we can be reasonably confident that our findings represent what *could* occur in "reality".

Nevertheless, there are a few threats to the validity of our findings that deserve mention. If only low-risk individuals ultimately comply with screening our results would have overstated the benefits of CRC screening in terms of reducing CRC incidence and mortality. However, the adenoma and CRC prevalence data that were used to inform the model were pooled from studies of average screening populations. As such, our findings likely already reflect that lower risk individuals are more likely than higher risk individuals to adhere with CRC screening recommendations [62]. In addition, adenoma and CRC progression rates across the classical adenoma-carcinoma sequence (Figure 1) are not known with certainty. A non-advanced adenoma could transition through to an advanced adenoma and onto a cancer more rapidly than we modeled. In this situation missing a lesion during a screening episode or as a result of non-adherence would lead to an increase in the incidence of CRC. This would reduce the effectiveness of screening. In contrast, slower progression rates would increase the potential of screening to identify lesions and reduce the incidence of CRC. If this were the case, CRC screening would be more effective than we found. Not all adenomas progress and indeed some data suggest that polyps can actually regress with time [63, 64]. On the other hand, a serrated polyp pathway has also been established and is characterized by lesions that are relatively flat, easy to miss and that progress rapidly to invasive cancer [65]. We adopted progression rates that were similar to those used in other modeling studies and made minor adjustments so that the number of cancers in our natural history model closely approximated the number of cancers observed in the control arms of the FOBT screening RCTs (please see the model validation section in the methods of the economic evaluation manuscript for more

details). We felt this strategy was the best option in light of the uncertainty surrounding this issue and varied our progression rate estimates in sensitivity analyses.

As already stressed, adherence is of critical importance to the main objective of CRC screening which is to reduce the incidence of and mortality from CRC. Yet screening rates continue to be low, especially in Canada where they have remained below 30% [56, 66]. In contrast, screening rates among non-minority insured individuals in the United States is estimated at nearly 60% [67]. Thus, there is a wide range of screening uptake in clinical practice, but the observed rates continue to lag behind targets that have generally been set at 70% [67]. Furthermore, it seems plausible that compliance rates might differ between strategies, but little is known regarding differential adherence rates. Nevertheless, although limited by available data our results were fairly robust to changes in both initial screening and surveillance uptake and a scenario of unequal initial adherence across the different screening tests.

It takes time for advancements in medical and surgical therapy for a disease such as CRC to translate into observed improvements in clinical outcome. Our stage-based CRC survival rates predate the widespread use of novel chemotherapies such as Bevacizumab. It is also becoming more common to surgically remove liver metastases in appropriate patients with advanced-stage disease. Whether these and future changes in the management of CRC translate into actual improvements in survival outside of clinical trials needs to be ascertained and could then be incorporated into our model. Of course the benefits of screening in reducing CRC-related mortality would be less if the disease were less deadly and thus screening might not have been as economically efficient had we significantly underestimated the effectiveness of the treatments for CRC.

Our conclusions were sensitive to our cost estimates including the cost of treating CRC and administering FIT. However, as previously discussed a scenario in which the currently funded biologic chemotherapies are excluded seems highly unlikely. Instead, it is predicted that the costs of managing patients with late-stage CRC will continue to climb. In addition, the dominance of FIT over no screening was sensitive to the level of additional administrative costs that might be associated with an annual screening program compared to one that recurs at a much less frequent interval (e.g. colonoscopy). FIT remained dominant over no screening as long as these potential additional administrative costs were no more than $10 per test. However, even at $30 per test, FIT remained the preferred screening strategy and was very attractive with a cost per QALY of $3,120 compared with no screening. Only when the additional administrative costs of FIT were $50 or more per test did colonoscopy have a lower cost per QALY than FIT. Since the completion of our work, demand and cost projections have been created for the proposed ACRCSP, which was mentioned above. According to the ACRCSP business plan where it was assumed that two-thirds of the 1.1 million age-eligible individuals would adhere with biennial FIT it was estimated that about 250,000 FIT tests would be completed annually across the Province. For this, an annualized operating budget of about 6.7 million dollars was requested (excluding capital costs), which translates into a cost of just under $27 per test. However, this estimate doesn't reflect the extra administrative costs above what would be required for a primary colonoscopy based program, and thus if anything is high. Refined estimates of the administrative costs of screening with FIT will need to be determined, but it seems unlikely that the magnitude of the administrative costs would threaten the validity of our results.

Finally, clinical practice guidelines make recommendations according to available evidence and prevailing expert opinion. Guideline-based rescreening and surveillance intervals not only influence the effectiveness of a given screening strategy, but the frequency of testing also has important implications on costs. It may be that other more or less frequent intervals are preferred from an economic standpoint. For instance, it is possible that adjustments in the age to start (e.g. 55 instead of 50) or end CRC screening and surveillance might offer better value for money. For instance, it is possible that an average risk 60 year-old individual with 10 years of negative FIT studies should stop screening, though clinical evidence to inform these decisions is lacking. As future evidence evolves, it will be important to re-visit the cost-effectiveness of CRC screening. In the meantime, it is possible to estimate the effects of altering screening and surveillance recommendations by adapting models such as ours.

4.3 Future Directions

Although several knowledge gaps remain, there are a few priority items that our model could help address in the near term. The following analyses would be of particular interest to health policy decision makers in Alberta, particularly those involved with the proposed Alberta Colorectal Cancer Screening Program:

1. It would be appropriate to re-run our analysis using Alberta-specific data. Nearly 40,000 colonoscopies have now been completed at the CCSC in Calgary. Thus, precise prevalence estimates could be generated and data on the rates and costs of colonoscopy complications including those unrelated to

bleeding and perforation could also be determined. Our model should be analyzed using not only the average risk prevalence rates from the CCSC, but above average risk prevalence rates as well. In addition to confirming the cost-effectiveness of FIT for CRC screening in Alberta, we could provide more precise estimates regarding the expected number of cancers and deaths from CRC in the province and the resource implications (e.g. demand for colonoscopy) associated with screening. Furthermore, this exercise would provide an opportunity to externally validate our model against population-level CRC incidence and mortality data in Alberta.

2. At the present time the rates of CRC mortality according to stage at diagnosis are not known in Alberta. However, efforts are currently underway to determine stage-based mortality rates for CRC. As discussed in the limitations section above, it will be important to confirm our findings in the context of current CRC survival data in Alberta.
3. The ACRCSP is still in the planning stages. Thus, it would be timely to provide additional information to the planning committee on the expected effects of delaying the onset of screening or potentially stopping screening among those with persistently negative tests as previously discussed. Adjustments such as these have the potential to reduce resource requirements and costs without negatively impacting clinical outcomes.
4. The FIT cut-off level recommended for referral to colonoscopy has generally been 75-150 ng/ml. One of the benefits of FIT is the ability to change the test-positive cut-off value. However, lowering the cut-off to increase sensitivity

comes at the cost of reducing specificity resulting in an increased rate of referral for unnecessary colonoscopies. Interestingly, a recent economic evaluation of FIT suggests that a cut-off of 50 mg/ml is preferred [32]. It will be critical to determine the preferred cut-off valve in Alberta before the ACRCSP is rolled-out as this will have important implications on the number of polyps and cancers missed, the number of colonoscopies demanded, and ultimately the costs of the Program. New evidence on the test performance of FIT at various cut-off values is becoming available and our model could be used to evaluate the impact of different set points on the cost-effectiveness of the ACRCSP.

4.4 Concluding Remarks

In conclusion, we have determined following a full-scale economic evaluation based on a validated Markov model populated with current evidence including a systematic review of polyp and CRC prevalence that average risk individuals should undergo CRC screening using FIT. The finding that CRC screening with FIT is cost saving is relatively novel. At a minimum this approach offers good value for health care money compared to other currently funded health interventions.

Knowledge translation is a critical component of research. It would have been suboptimal had our findings gone unnoticed and ultimately buried in the medical literature. However, at the time of the writing of this thesis the economic evaluation manuscript has received over 6500 online viewings. This is very encouraging. Even more important is the fact the ACRCSP under development has been informed by this work

and is centered on FIT for average risk screening. Finally, our model should be seen as a valuable resource to help guide and refine health policy decision making regarding CRC screening today and in the future.

APPENDIX 1: EFFECTIVENESS DATA FOR THE DIFFERENT CRC SCREENING OPTIONS

Fecal Occult Blood Test

Landmark randomized controlled trials in the 1990s [68-70] determined that regular (annual and biennial) testing for occult blood in the stool reduced CRC mortality by between 15% [68] and 33% [70].

The traditional fecal occult blood tests (FOBTs) rely on an oxidative reaction created from the presence of pseudoperoxidase activity found in hemoglobin. The reaction changes guaiac, a colorless compound blue. Test performance characteristics for different FOBT tests vary widely. Factors that can affect test performance include fecal hydration (hydration increases sensitivity while reducing specificity) and foods containing pseudoperoxidase or peroxidase activity (e.g. red meat, broccoli, and cauliflower) that can falsely turn the indicator dye blue [71].

A single "test" is defined as 2 samples collected from 3 discrete bowel movements. Non-rehydrated samples are generally preferred to limit false positives [72]. Studies in which both FOBT and a reference standard colonoscopy were performed in all individuals have shown FOBT to have a sensitivity for CRC ranging from 13% [38] to 41% [73] for a single test. Although single test sensitivities for CRC and advanced polyps are poor, FOBT test performance increases with repeated testing given multiple detection opportunities [74]. Advantages of FOBT include its widespread availability, relative ease of use, safety and low cost.

Fecal Immunochemical Test

Fecal immunochemical tests (FIT) use antibodies directed against human globin to detect blood in stool. Compared with the qualitative FOBT test (positive/negative), FIT tests are quantitative and different thresholds can be set which determine their level of sensitivity and specificity [75]. Although not widely used at present in North America, studies have shown FIT to have superior test performance over FOBT [42]. Furthermore, a single test only requires sampling from one bowel movement and there is no need for dietary restrictions, both of which may account for the observed improvement in patient participation rates compared to FOBT [42]. FIT is promising for population-based CRC screening.

Fecal DNA

The fecal DNA test detects DNA markers shed from polyps and CRC into stool. The test is not reliant on the presence of bleeding, which can be intermittent and even absent altogether for CRC. Like the guaiac- and immunochemical-based stool tests, fecal DNA offers a non-invasive option for patients. Although an earlier study showed it to have relatively low sensitivity for advanced neoplasia (advanced adenomas and CRC), it was still better than Hemoccult II (a guaiac-based FOBT) at 18% compared to 11%, [38]. However, using a newer generation of the test (SDT-2) in a screening setting Ahlquist et al. [76] showed that fecal DNA had a sensitivity of 46% compared to 16% for Hemoccult II for identifying advanced neoplasia. The enhanced sensitivity of the newer generation test was most apparent for adenomas. The major drawback of fecal DNA remains its high

cost. US modeling studies to date have not shown it to be cost-effective compared to other CRC screening strategies [48, 50].

Sigmoidoscopy and Colonoscopy

Sigmoidoscopy and colonoscopy involve insertion of a flexible endoscope via the anus into the colon following a purging laxative. The endoscope is inserted to the splenic flexure during sigmoidoscopy and to the cecum during colonoscopy. It is recommended that those found to have an adenomatous polyp on sigmoidoscopy undergo a full colonoscopy to search for associated proximal pathology. In addition to high sensitivity, the major advantage of sigmoidoscopy and colonoscopy is the ability to remove polyps and obtain biopsies during the same procedure. However, accompanying this therapeutic potential comes the added risk of bleeding, perforation and very rarely death [72].

Colonoscopy has been considered by many to be the most effective screening modality [23]. Epidemiological evidence has suggested 70-90% reductions in CRC incidence following mass screening with colonoscopy [15]. However, the magnitude of the benefit of colonoscopy over sigmoidoscopy has been called into question given that a recent case-control study by Baxter et al. [41] showed a mortality reduction for left, but not right sided CRC using colonoscopy. The reason for this finding is not entirely clear, but may be due to suboptimal colonic preparation limiting visualization, a higher prevalence of flat or “sessile” neoplasia or potentially a different tumor biology in the right colon [41].

The widely anticipated results of randomized trials using flexible sigmoidoscopy are beginning to emerge. Flexible sigmoidoscopy with or without a single round of FOBT testing compared to no screening did not show a reduction in CRC mortality in the

intention to screen analysis [77]. However, among those who were compliant with screening (per protocol analysis) there appeared to be a 59% reduction in CRC mortality at any location and a 76% reduction in the rectosigmoid (left-sided) area [77]. A larger randomized study of once only flexible sigmoidoscopy versus no screening between the ages of 55 and 64 showed a 23% reduction in the incidence of CRC and a 31% reduction in mortality [78].

Barium Enema and Computed Tomographic Colonography

Until recently barium enema was the only radiographic modality capable of evaluating the entire colon. Although minimally invasive, a full colonic preparation is still required. It has only moderate test performance at best [79] and a colonoscopy is required following a positive exam to confirm and remove polyps. Furthermore, even though it is relatively safe, patients often find it uncomfortable [80]. Barium enema is used infrequently for screening in Alberta [40].

Computed tomographic colonography (CTC) or virtual colonoscopy represents a major advance over barium enema. Recent studies have demonstrated test performance characteristics for polyps >10 mm and CRC that rival those of colonoscopy [39, 81, 82]. However, like barium enema it is only diagnostic and issues including polyp size thresholds for reporting, re-screening intervals, extra-colonic findings and radiation exposure persist [83]. Furthermore, the test is expensive and there is uncertainty regarding its cost-effectiveness. While some mathematical modeling studies that permit small polyps to be ignored suggest that CTC might be cost-effective [44], others have not shown it to be cost-effective compared to the commonly available CRC screening

modalities [45-47].

APPENDIX 2: ELEMENTS OF AN ECONOMIC EVALUATION

The following is a summary of the elements that should be included in an economic evaluation. Further detail regarding each of these elements can be found in the guidelines for the economic evaluation of health technologies published by the Canadian Agency for Drugs and Technologies in Health (CADTH) [33].

Study Question

A well-defined and focused study question should be formalized prior to undertaking an economic evaluation. The study question should describe the patient population, the intervention and relevant comparators and define the perspective of the analysis. Finally, in framing the study question one must ensure that the results generated from the analysis are relevant to the target audience for which the research was intended.

Target Population

The economic impact of a program or intervention depends on the target population being evaluated. Target populations may be defined based on baseline demographic characteristics (e.g., age and sex) or according to the presence or absence of specific conditions. In addition, populations can be defined by setting (e.g., community or hospital), geographic location, or usual patterns of treatment. Although the evaluation should analyze the entire population defined in the study question, it may be appropriate to conduct stratified analyses if there is heterogeneity in important variables within the target population.

Type of Economic Evaluation

Selecting the appropriate type of evaluation depends on the research question and the availability of outcome data. Analysts should justify the chosen type of evaluation. CADTH recommends using a cost-utility analysis as the primary analysis [33].

Comparators

A comparator is any alternative that has the potential to be replaced by the intervention in question. The selection of appropriate comparators is critically important as this choice including a decision to omit a given comparator can directly impact the results of the analysis and the relevance to the target audience. In general, all reasonable alternatives should be considered even though not all necessarily have to be selected for the analysis. However, a decision to omit a given comparator should be justified. This process can be particularly challenging when multiple comparators are available and variations in clinical practice exist.

The reference case should be “usual care,” which is generally the most common or widely used comparator. The status quo may be a “do nothing” approach. It is also advisable to consider future comparators that have the potential to be adopted.

Perspective

The perspective of the evaluation should be based on the intended target audience. Examples include the public payer perspective, the publicly funded health care system perspective and the societal perspective. The differences between these relate to the costs that are considered relevant. CADTH recommends that the publicly funded health care

system be chosen for the primary analysis [33]. The costs considered relevant to the publicly funded health care system include both direct costs (e.g. hospital services, physician billings, drug costs, etc.) and indirect or non-medical costs. Non-medical costs are patient and caregiver time and travel costs. Productively costs (e.g. lost wages due to time off work, the cost incurred by an employer to hire new staff) are also considered in addition to non-medical costs in the wider societal perspective. The costs associated with adopting a wider perspective should be reported separately where it is likely that they will impact the results of the analysis.

Effectiveness

Whereas efficacy refers to how well an intervention performs in a controlled setting (i.e. within the confines of an RCT), effectiveness refers to how well the intervention performs under "real life" conditions. In reality, a technology is often used by a variety of end users with different skill sets on individuals who are less informed, less compliant and have co-morbid conditions that would have excluded them from an RCT. As one might expect, "real life" patients tend to respond less favourably to treatments than participants in RCTs [84].

Decision-makers are primarily concerned with the effectiveness rather than the efficacy of an intervention when making decisions regarding whether or not to fund. Thus, it is recommended that the outcomes and costs informing an economic evaluation be based on effectiveness rather than efficacy data when possible. This can be difficult when evaluating non-drug technologies, which tend to get approved following less rigorous clinical testing than medications. Finally, the impact of adverse events on the

intervention should be accounted for in the analysis as they may affect patient compliance, morbidity and mortality, quality of life and resource use. Deviations from these practices should be justified.

Efficacy data are frequently limited and tend to surface after a technology has been implemented. Factors that have the potential to modify effectiveness should be identified such as patient adherence with treatment and the knowledge and skill of providers, both of which can influence outcomes. Once identified, these variables can be varied in sensitivity analyses in order to model their effect on "real life" practice.

Time Horizon

The time horizon of an economic evaluation refers to the duration of time over which costs and outcomes are considered in the analysis. It should be long enough to capture all the relevant potential differences in these variables between the intervention and comparators. Analysts are encouraged to consider a lifetime time horizon in the primary analysis, especially for chronic conditions [33]. However, it may be appropriate to consider a short-term analysis as long as the decision is justified [45].

Valuing Outcomes: Cost-Utility Analysis

The preferred outcome measure for a cost-utility analysis is the QALY. Please see section 1.2.5 for general information on quality of life, utilities and the derivation of a QALY.

There is controversy whether utilities should be derived from patients with current or prior exposure to a given health state or from a representative sample of the general

public. It is known that preference valuation varies according to one's experience with a given condition [85]. The indirect measurement instruments discussed in section 1.2.5 are based on surveys of preferences from the general public and it is preferred that direct utility measurement likewise be obtained from a sample of the general public who have been adequately informed about the health states being valued. The justification for this approach is that the general public are the payers of the publicly funded health care system and represent potential patients with the condition(s).

Resource Use and Costs

The relevant resource items to consider along with their respective costs are determined by the perspective taken in the analysis (see above). Both present and future costs resulting from the intervention and alternatives are potentially relevant.

There are different categories of health care costs. Direct health care costs include those resulting from hospital services, laboratory tests and physician fees. Indirect costs, also known as non-medical costs, include 1) direct costs to patients and their families for receiving care (e.g. travel costs), 2) time costs to patients and their families while receiving care and 3) costs resulting from lost productivity. It remains controversial whether or not and how to account for lost productivity costs [86]. However, it is now generally recommended that the first two categories of non-medical costs be included in economic evaluations of health care programs, at least when using the perspective of the publicly-funded health care system [33].

Two methods for costing exist [87]. Gross costing considers items on an aggregate basis such as the average cost per hospital day. In contrast, micro-costing considers each

individual cost component, an approach which can produce precise item estimates, but is time consuming. In reality a combination of these approaches is used and items that have the potential to drive the results should be quantified more precisely. Uncertainly in important cost estimates should be disclosed and sensitivity analyses to determine the impact of these uncertain cost assumptions should be performed.

Resource use and cost data can be obtained from a number of sources including clinical trials, administrative data and the published literature. Using costs from other countries in the Canadian setting may not be appropriate. Lastly, costs may only be available for a previous time period and should be updated accordingly (e.g. inflation adjusted often using the Consumer Price Index).

Discounting

Individuals prefer resources now to waiting for them in the future. This notion of time preference is underscored by the existence of interest rates. An individual demands interest from a bank in order to save or postpone consumption. Indeed, the rate of interest reflects the opportunity cost of money [29]. However, the reason for time preference involves more than just interest rates. Some individuals simply have a short-term view of life and most would agree that the future is uncertain. In addition, given positive economic growth many individuals expect to be wealthier in the future and hence a dollar today is worth more than one in the future.

In order to perform an economic evaluation today involving costs and health outcomes in the future, an analyst must compute the value of future resource transactions in today's dollars. The present value, or value today, of a future dollar is determined by

how far in the future it is obtained and on the rate at which it is discounted. A dollar invested today at an interest rate of 5% will generate $1.05 a year from now ($1.00 X 1.05 = $1.05). The calculation is reversed to determine the present value of a dollar that we expect to obtain a year from now (present value X 1.05 = $1.00 where present value = 0.95). The general formula for present value is:

$$PV = \frac{FV}{(1+r)t}$$

where PV is present value, FV is future value, r is the discount rate (expressed as a decimal fraction and t is time. Using this formula, the present value of $100 attained 10 years from now, discounted at 5% annually is $PV = \$100/(1 + 0.05)^{10} = \$61$.

It is generally accepted that health outcomes be discounted in an identical manner to that described above. Although controversy exists, the basic argument for discounting future health outcomes is that they are being valued relative to dollars that could have yielded even more dollars in the future if not spent today to generate the outcomes of interest [29]. In Canada, the recommended discount rate is 5% [33]. However, it should also be set to 0% in a sensitivity analysis to assess the impact of discounting.

Variability and Uncertainty

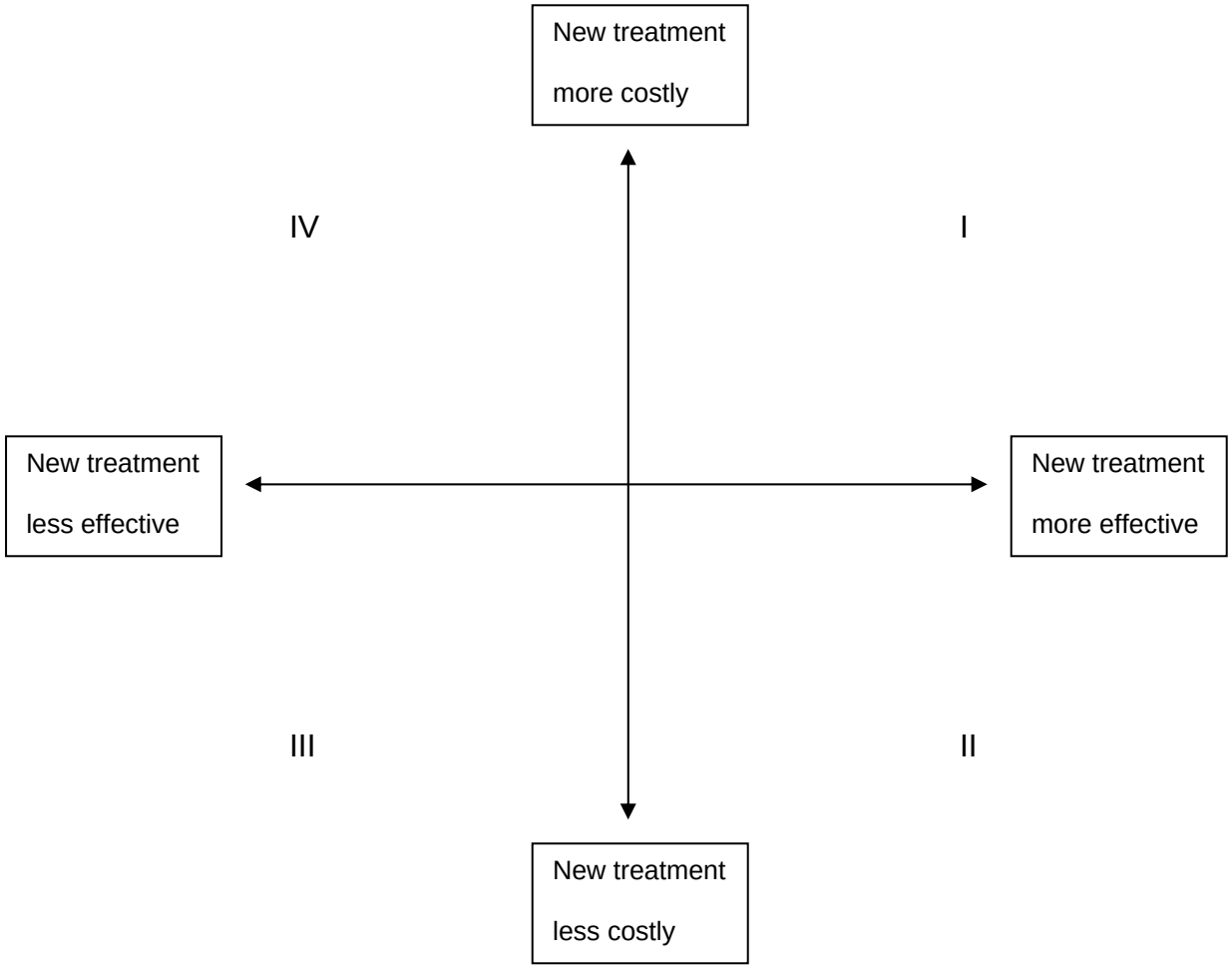
Variability reflects differences in the values of parameters that result from changes in circumstances. For example, the effectiveness of an intervention may differ between older and younger individuals. In modeling variability is handled primarily

through scenario analyses, which involve simultaneously changing all the relevant model parameters unique to the scenario [33].

Uncertainty occurs when the true value of a parameter is unknown. In reality the "true" value of any measured entity is always unknown because measurement is imperfect. In the context of an economic evaluation uncertainty can be divided into parameter uncertainty (costs, effects, etc.) and model uncertainty (structure, assumptions, etc.). Parameter uncertainty refers to uncertainty in a probabilistic sense; i.e. uncertainty due to random chance from sampling. Parameter uncertainty can be dealt with through deterministic and probabilistic sensitivity analyses. In a deterministic sensitivity analysis parameters are analyzed as point estimates which are then varied over plausible ranges. In a probabilistic sensitivity analysis inputs are analyzed as probability distributions to more accurately reflect their full uncertainty. A probabilistic sensitivity analysis is preferred when possible as it can provide a more complete assessment of the uncertainty associated with all of the inputs in the model simultaneously [33].

Model uncertainty is created through the choices and assumptions made by the modeller. Examples include simplifications in the underlying biology of a disease process and the time horizon or discount rate chosen. Model uncertainty can be dealt with through deterministic sensitivity analyses using alternative model assumptions. It is of paramount importance to give due diligence to uncertainty at every level of a model. This is required so that the reader is capable of judging whether the results are meaningful and robust.

APPENDIX 3: COST-EFFECTIVENESS PLANE



Adapted from Drummond [1]

APPENDIX 4: MODEL VALIDATION

Ensuring the internal validity of our model was a major focus of this thesis. Interval validity was approached in two stages. In the first stage the face validity of the model was established. For this clinical experts (Drs. Heitman and Hilsden) designed and agreed on the model structure and flow to simulate "real life" as closely as possible. In designing the model a number of assumptions were made. Firstly, all polyps are considered potentially important. Cancer formation through the classic adenoma carcinoma sequence is assumed with CRC arising only from large ($\geq$ 10 mm) adenomatous polyps. Moreover, those with a previous history of adenomatous polyps or cancer are assumed to have a more rapid adenoma-carcinoma sequence. It is also assumed that missed cancers become symptomatic within a five-year interval. Once CRC is discovered through screening or clinically manifest, mortality is modeled according to cancer stage. Lastly, the effectiveness of CRC screening in terms of reducing mortality in the model is realized through a more favourable cancer stage distribution and through removal of polyps.

After the model was created, extensive "debugging" exercises were undertaken to confirm its technical accuracy. We assessed for logical inconsistences by evaluating our model under hypothetical extreme value scenarios (e.g. 0% and 100% adherence) to ensure that the results made sense. This permitted us to check for and correct all syntactical errors.

In the second stage of our validation we compared our "no screening" and FOBT strategies with the results of the randomized trials that compared FOBT with "no

screening" [68-70]. In this exercise, we also performed a "between-model validation" by comparing our model results to those of an independently developed model that was created for the US Multi-Society Task Force [88]. In order to simulate the results of the control arms of the clinical trials, we altered our time horizon to be consistent with the RCT follow-up periods and performed first order Monte Carlo simulation on the same number of simulated patients that were studied in the trials. Table 1a of this appendix found below illustrates the results comparing our no screening/natural history arm to that of the control arms of the FOBT RCTs and the no screening strategy of the Multi-Society Task Force model in terms of the total number of cancers. Our results were nearly identical to two of the RCTs and the Multi-Society Task Force Model and were within 10% of the third RCT, suggesting that our model was well calibrated to modelling the progression of CRC.

The study by Mandel et al. screened patients with FOBT annually and our model evaluated annual screening using the stool tests in the base case. As such, we validated our FOBT strategy by comparing our results to those of Mandel et al. and also compared our results to the model created for the Multi-Society Task Force. In a similar fashion to that described above we altered our time horizon, sample size and assumed similar adherence and test performance characteristics described in the RCT. We also did this for the between-model validation. Once again, Table 1b shows that the total number of cancers predicted by our model closely matched that of the RCT and the other model, suggesting that our model was well calibrated to evaluating the impact of fecal based screening in CRC.

The only difference between the FOBT strategy and the other stool based strategies (FIT and fecal DNA) involved test performance. Thus, the validated FOBT strategy was simply modified for the other fecal-based screening strategies.

We validated our colonoscopy strategy by first assuming that it was a perfect test (sensitivity and specificity 100%) and that it would be performed annually with 100% compliance. With this, all the CRCs were found on initial screening and no CRCs occurred subsequently (because polypectomies were performed on all patients who developed polyps). To validate the CTC strategy, we re-analyzed it assuming that CTC had similar test performance characteristics as colonoscopy for polyps and cancers, that adherence with colonoscopy after a positive CTC was 100%, and that colonoscopy after a positive CTC was a perfect test. Using these assumptions, we noted that the number of missed cancers occurring in the CTC strategy was nearly equivalent to the number of missed cancers occurring in the colonoscopy strategy.

Having completed these exercises, we were confident that our model had face validity, was internally valid, that it was calibrated to other data sources, and that it compared favourably with other models (between-model validation).

Table 1a: No Screening/Natural History Validation

Source	Number of Patients	Years of Follow-up	Number of CRCs
Mandel et al. RCT	15,394	13	356
Our model with Mandel et al. data	15,394	13	354
Kronborg et al. RCT	30,966	10	483
Our model with Kronborg et al. data	30,966	10	494
Hardcastle et al. RCT	76,384	8	856
Our model with Hardcastle et al. data	76,384	8	945
US Multi Society Task Force Model	100,000	Lifetime	4,988
Our model with Multi-Society Task Force assumptions	100,000	Lifetime	4,929

Table 1b: FOBT Screening Strategy Validation

	Mandel et al. FOBT arm	Our model: Mandel data	Multi-Society Task Force model	Our model: Multi-Society assumptions
Number of patients	15,570	15,570	100,000	100,000
Years Follow-up	13	13	Lifetime	Lifetime
Initial screening compliance (%)	85	85	100	100
Re-screening compliance (%)	75	75	100	100
Sensitivity of FOBT for cancer	25	25	60	60
Total CRC	323	307	2,610	2,702

REFERENCES:

1. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 2005.

2. Heitman SJ, Ronksley PE, Hilsden RJ, et al. Prevalence of adenomas and colorectal cancer in average risk individuals: A systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009.

3. Heitman SJ, Hilsden RJ, Au F, et al. Colorectal cancer screening for average-risk North Americans: an economic evaluation. PLoS Med 2010;7:e1000370.

4. Canadian Cancer Society NCIoC. Canadian Cancer Statistics, 2007.

5. 2011 colorectal cancer statistics. Toronto: Colorectal Cancer Association of Canada, 2011.

6. Fuchs CS, Giovannucci EL, Colditz GA, et al. A prospective study of family history and the risk of colorectal cancer. N Engl J Med 1994;331:1669-74.

7. Nguyen SP, Bent S, Chen YH, et al. Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009;7:676-81 e1-3.

8. Botteri E, Iodice S, Bagnardi V, et al. Smoking and colorectal cancer: a meta-analysis. Jama 2008;300:2765-78.

9. Renehan AG, Tyson M, Egger M, et al. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008;371:569-78.

10. Maroun J, Ng E, Berthelot JM, et al. Lifetime costs of colon and rectal cancer management in Canada. Chronic Dis Can 2003;24:91-101.

11. Yabroff KR, Mariotto AB, Feuer E, et al. Projections of the costs associated with colorectal cancer care in the United States, 2000-2020. Health Econ 2008;17:947-59.

12. Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med 2004;350:2335-42.

13. Drucker A, Skedgel C, Virik K, et al. The cost burden of trastuzumab and bevacizumab therapy for solid tumours in Canada. Curr Oncol 2008;15:136-42.

14. Bradley CJ, Yabroff KR, Dahman B, et al. Productivity costs of cancer mortality in the United States: 2000-2020. J Natl Cancer Inst 2008;100:1763-70.

15. 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;329:1977-81.

16. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer 1975;36:2251-70.

17. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-32.

18. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-67.

19. Cummings OW. Pathology of the adenoma-carcinoma sequence: from aberrant crypt focus to invasive carcinoma. Semin Gastrointest Dis 2000;11:229-37.

20. Leddin DJ, Enns R, Hilsden R, et al. Canadian Association of Gastroenterology position statement on screening individuals at average risk for developing colorectal cancer: 2010. Can J Gastroenterol 2010;24:705-14.

21. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008;149:627-37.

22. 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. Gastroenterology 2008;134:1570-95.

23. Rex DK, Johnson DA, Anderson JC, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2008. Am J Gastroenterol 2009;104:739-50.

24. Rabeneck L, Paszat LF. A population-based estimate of the extent of colorectal cancer screening in Ontario. Am J Gastroenterol 2004;99:1141-4.

25. Schrag D. The price tag on progress--chemotherapy for colorectal cancer. N Engl J Med 2004;351:317-9.

26. Ergun H, Gulmez SE, Tulunay FC. Cost-minimization analysis comparing topiramate with standard treatments in migraine prophylaxis. Eur Neurol 2007;58:215-7.

27. Mark DB, Hlatky MA, Califf RM, et al. Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. N Engl J Med 1995;332:1418-24.

28. Khanna D, Tsevat J. Health-related quality of life--an introduction. Am J Manag Care 2007;13 Suppl 9:S218-23.

29. Hunink M, Glasziou P. Decision making in health and medicine. Integrating evidence and values. Cambridge: Cambridge University Press, 2001.

30. Dhalla I, Laupacis A. Moving from opacity to transparency in pharmaceutical policy. Cmaj 2008;178:428-31.

31. Flanagan WM, Le Petit C, Berthelot JM, et al. Potential impact of population-based colorectal cancer screening in Canada. Chronic Dis Can 2003;24:81-8.

32. Wilschut JA, Hol L, Dekker E, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology 2011;141:1648-55 e1.

33. Guidelines for the economic evaluation of health technologies. In: Health CAfDaTi, ed. Third Edition ed. Ottawa, Canada, 2006.

34. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. Value Health 2003;6:9-17.

35. Laupacis A, Feeny D, Detsky AS, et al. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. Cmaj 1992;146:473-81.

36. National Commitee on Colorecal Cancer Screening. Technical report for the National Committee on Colorectal Cancer Screening, 2003.

37. Pignone M, Saha S, Hoerger T, et al. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:96-104.

38. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;351:2704-14.

39. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003;349:2191-200.

40. McGregor SE, Hilsden RJ, Murray A, et al. Colorectal cancer screening: practices and opinions of primary care physicians. Prev Med 2004;39:279-85.

41. Baxter NN, Goldwasser MA, Paszat LF, et al. Association of colonoscopy and death from colorectal cancer. Ann Intern Med 2009;150:1-8.

42. 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;135:82-90.

43. Hassan C, Pickhardt P, Laghi A, et al. Computed tomographic colonography to screen for colorectal cancer, extracolonic cancer, and aortic aneurysm: model simulation with cost-effectiveness analysis. Arch Intern Med 2008;168:696-705.

44. Pickhardt PJ, Hassan C, Laghi A, et al. Cost-effectiveness of colorectal cancer screening with computed tomography colonography: the impact of not reporting diminutive lesions. Cancer 2007;109:2213-21.

45. Heitman SJ, Manns BJ, Hilsden RJ, et al. Cost-effectiveness of computerized tomographic colonography versus colonoscopy for colorectal cancer screening. Cmaj 2005;173:877-81.

46. Ladabaum U, Song K, Fendrick AM. Colorectal neoplasia screening with virtual colonoscopy: when, at what cost, and with what national impact? Clin Gastroenterol Hepatol 2004;2:554-63.

47. Vijan S, Hwang I, Inadomi J, et al. The cost-effectiveness of CT colonography in screening for colorectal neoplasia. Am J Gastroenterol 2007;102:380-90.

48. Song K, Fendrick AM, Ladabaum U. Fecal DNA testing compared with conventional colorectal cancer screening methods: a decision analysis. Gastroenterology 2004;126:1270-9.

49. Itzkowitz SH, Jandorf L, Brand R, et al. Improved fecal DNA test for colorectal cancer screening. Clin Gastroenterol Hepatol 2007;5:111-7.

50. Parekh M, Fendrick AM, Ladabaum U. As tests evolve and costs of cancer care rise: reappraising stool-based screening for colorectal neoplasia. Aliment Pharmacol Ther 2008;27:697-712.

51. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med 2008;149:659-69.

52. Ho C, Heitman S, Membe SK, et al. Computed tomographic colonography for colorectal cancer screening in an average risk population: Systematic review and economic evaulation. In: Health CAfDaTi, ed, 2008.

53. Chu E. An update on the current and emerging targeted agents in metastatic colorectal cancer. Clin Colorectal Cancer 2012;11:1-13.

54. Cancer Care Ontario. Colorectal Cancer Screening, 2011.

55. McGregor S, Hilsden R, Yang H. Physician barriers to population-based, fecal occult blood test-based colorectal cancer screening programs for average-risk patients. Can J Gastroenterol 2010;24:359-64.

56. McGregor SE, Hilsden RJ, Li FX, et al. Low uptake of colorectal cancer screening 3 yr after release of national recommendations for screening. Am J Gastroenterol 2007;102:1727-35.

57. Ritvo P, Myers R, Del Giudice ME, et al. Fecal occult blood testing: people in Ontario are unaware of it and not ready for it. Can Fam Physician 2009;55:176-177 e4.

58. Heitman SJ, Au F, Manns BJ, et al. Nonmedical costs of colorectal cancer screening with the fecal occult blood test and colonoscopy. Clin Gastroenterol Hepatol 2008;6:912-917 e1.

59. Leddin D, Bridges RJ, Morgan DG, et al. Survey of access to gastroenterology in Canada: the SAGE wait times program. Can J Gastroenterol 2010;24:20-5.

60. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut 2012.

61. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Engl J Med 2012;366:697-706.

62. Laiyemo AO, Doubeni C, Pinsky PF, et al. Race and colorectal cancer disparities: health-care utilization vs different cancer susceptibilities. J Natl Cancer Inst 2010;102:538-46.

63. Hofstad B, Vatn MH, Andersen SN, et al. Growth of colorectal polyps: redetection and evaluation of unresected polyps for a period of three years. Gut 1996;39:449-56.

64. Hoff G, Foerster A, Vatn MH, et al. Epidemiology of polyps in the rectum and colon. Recovery and evaluation of unresected polyps 2 years after detection. Scand J Gastroenterol 1986;21:853-62.

65. Oono Y, Fu K, Nakamura H, et al. Progression of a sessile serrated adenoma to an early invasive cancer within 8 months. Dig Dis Sci 2009;54:906-9.

66. Zarychanski R, Chen Y, Bernstein CN, et al. Frequency of colorectal cancer screening and the impact of family physicians on screening behaviour. Cmaj 2007;177:593-7.

67. Klabunde CN, Brown M, Ballard-Barbash R, et al. Cancer Screening - United States, 2010: Centers for Disease Control and Prevention, 2012.

68. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472-7.

69. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467-71.

70. 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;328:1365-71.

71. Ouyang DL, Chen JJ, Getzenberg RH, et al. Noninvasive testing for colorectal cancer: a review. Am J Gastroenterol 2005;100:1393-403.

72. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence. Gastroenterology 2003;124:544-60.

73. Collins JF, Lieberman DA, Durbin TE, et al. Accuracy of screening for fecal occult blood on a single stool sample obtained by digital rectal examination: a comparison with recommended sampling practice. Ann Intern Med 2005;142:81-5.

74. Church TR, Ederer F, Mandel JS. Fecal occult blood screening in the Minnesota study: sensitivity of the screening test. J Natl Cancer Inst 1997;89:1440-8.

75. Hol L, Wilschut JA, van Ballegooijen M, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer 2009;100:1103-10.

76. Ahlquist DA, Sargent DJ, Loprinzi CL, et al. Stool DNA and occult blood testing for screen detection of colorectal neoplasia. Ann Intern Med 2008;149:441-50, W81.

77. Hoff G, Grotmol T, Skovlund E, et al. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. Bmj 2009;338:b1846.

78. 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.

79. Rockey DC, Paulson E, Niedzwiecki D, et al. Analysis of air contrast barium enema, computed tomographic colonography, and colonoscopy: prospective comparison. Lancet 2005;365:305-11.

80. Van Ness MM, Chobanian SJ, Winters C, Jr., et al. A study of patient acceptance of double-contrast barium enema and colonoscopy. Which procedure is preferred by patients? Arch Intern Med 1987;147:2175-6.

81. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008;359:1207-17.

82. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007;357:1403-12.

83. Whitlock EP, Lin JS, Liles E, et al. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2008;149:638-58.

84. Rawson NS. "Effectiveness" in the evaluation of new drugs: a misunderstood concept? Can J Clin Pharmacol 2001;8:61-2.

85. Brazier J, Deverill M, Green C, et al. A review of the use of health status measures in economic evaluation. Health Technol Assess 1999;3:i-iv, 1-164.

86. Oostenbrink JB, Koopmanschap MA, Rutten FF. Standardisation of costs: the Dutch Manual for Costing in economic evaluations. Pharmacoeconomics 2002;20:443-54.

87. Brouwer W, Rutten F, Koopmanschap M. Costing in economic evaluations. In: Drummond M, McGuire A, eds. Economic evaluation in health care: merging theory with practice. Oxford: Oxford University Press, 2001:68-93.

88. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 1997;112:594-642.