A microsimulation study of the benefits and costs of screening for colorectal cancer

Christopher Eric Stevenson

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

September 2000
Amended April 2001

Declaration

Except where it is otherwise indicated, the work in this thesis is my own, and is based on original research conducted at the National Centre for Epidemiology and Population health, The Australian National University.

Chris Stevenson

September 2000

This revised version of the thesis incorporates minor amendments as suggested by the examiners.

April 2001

Acknowledgments

I thank my panel of supervisors, Professor Susan Wilson, Dr Jim Butler, Dr Terry O'Neil and Dr Jane Matthews, for their guidance and valuable insights throughout the preparation of this thesis.

Academic staff and students at the National Centre for Epidemiology and Population Health were always helpful and supportive, and I am grateful to the Centre for the opportunity to spend a year studying there full time. However, most of this degree was completed as a part time student and during this time I especially valued the daily support of my colleagues at the Australian Institute of Health and Welfare--particularly my fellow students Edouard d'Espaignet and Stan Bennett and the Institute's librarian Judith Abercromby. The Institute also provided valuable support and encouragement particularly in the provision of study leave without which this thesis would have been impossible.

Most importantly, I thank my wife Mandy for her constant support during all phases of my study and my son Lachlan for making the whole thing worthwhile.

Abstract

This thesis examines the benefits and costs of screening for colorectal cancer in the context of an organised population screening programme. It uses microsimulation modelling to derive an optimally cost-effective screening protocol for various combinations of the available screening tests.

First a mathematical model for the natural history of colorectal cancer is derived, based on analyses of Australian population and hospital-based cancer registries combined with data from published studies. Then a model for population based screening is derived based mainly on data from published screening studies, including the four major published randomised controlled trials of faecal occult blood test (FOBT) screening. These two models are used to simulate the application of a screening programme to the Australian population. The simulations are applied to a period of 40 years following 1990 (the study's base year), with both costs and benefits discounted back to the base year at an annual rate of 3%.

The models are applied to simulating a population screening programme based on FOBT with a colonoscopy follow up of positive tests. This simulation suggests that the optimal application of such a programme would be to offer annual screening to people aged 50 to 84 years. Such a programme would lead to a cumulative fall in years of life lost to colorectal cancer (YLL) of 28.5% at a cost per year of life saved (YLS) of $8,987. These costs and benefits are consistent with those arising from other currently funded health interventions. They are also consistent with the cost per YLS which Australian governments appear willing to pay for health interventions when justified on the basis of cost-effectiveness. The fall in colorectal cancer deaths from this screening programme should be first detectable by a national monitoring system after around three years of screening. However the full benefits from screening would not be realised before around 30 years of screening.

These simulations are based on the standard guaiac FOBT, but the results suggest that significant cost-effective gains could be made by using the newer immunochemical FOBT. Further cost-effect gains could be made by offering sigmoidoscopy every five years in addition to annual FOBT.

The models are then applied to simulating population screening programmes using colonoscopy and sigmoidoscopy as primary screening tools. Offering colonoscopy every ten years to all people aged from 45 to 85 leads to an overall fall in cumulative YLL of 37.6%, at a cost of $15,585 per YLS. Offering sigmoidoscopy every three years to all people aged 40 to 85 leads to an overall fall in cumulative YLL of 29.1%, at a cost of $4,862 per YLS. Both of these cost and benefit results are also consistent with the cost per YLS which Australian governments appear willing to pay. The fall in deaths with colonoscopy screening would also be detectable after three years of screening but the fall with sigmoidoscopy screening would not be detectable until after six years of screening. Sigmoidoscopy would need around 35 years of

screening to reach its potential gains while colonoscopy screening would not reach its full potential during the 40 year screening period.

Finally the models are applied to targeting people at higher risk of cancer. The results show that offering colonoscopy every five years to people at higher risk because of a family history of colorectal cancer is a cost-effective addition to the annual FOBT screening programme.

An earlier version of chapter two of this thesis has been published as

Stevenson CE 1995. Statistical models for cancer screening. *Statistical Methods in Medical Research*; 4: 19–23.

An expanded version of chapter two, along with parts of chapter one, has been published as

Stevenson CE 1998. Models of screening. In: *Encyclopedia of Biostatistics.* Armitage P, Colton T, eds. John Wiley and Sons Ltd, pp 3999–4022.

Table of contents

1 Introduction

1.1 Study aims

The object of this thesis is to develop a mathematical model of screening for colorectal cancer which will (i) summarise the natural history of the cancer and (ii) enable an analysis of the effect of different screening strategies on its morbidity and mortality.

Previous studies have shown that screening for both breast and cervical cancer is more effective and cheaper when carried out in an organised programme (AHMAC 1990, AHMAC 1991). This is also likely to be the case with colorectal cancer screening. The advantages of an organised approach include:

- Agreed screening age ranges and intervals;
- Organised and targeted recruitment strategies, including, for example, personalised initial invitations to attend screening and personalised reminder letters for subsequent screens;
- Organised quality control in the storage, handling and interpretation of screening material such as FOBT slides;
- Organised systems to ensure people with positive test results are followed-up;
- Agreed protocols for appropriate management of people with minor screen-detected abnormalities such as small polyps;
- Organised monitoring of screening to provide
    - data to enable accurate assessment of regional, State-Territory and national screening rates, as well as screening rates in population subgroups which may be subject to significant under screening;
    - comprehensive monitoring of the accuracy of screening tests;
    - a continuing evaluation and monitoring of the screening programme activities in relation to agreed screening goals and targets.

My study will examine screening in the context of an organised population screening programme.

The aim of this work is to answer the following questions:

1. What are the likely benefits and costs of colorectal cancer screening in Australia?

Firstly I will investigate the costs and benefits of such a programme in terms of its effect on overall cancer mortality. As part of this I will also investigate whether government funding of such a programme is justifiable in cost-effectiveness terms. Secondly, I will investigate the benefits of such screening to screening participants. Finally I will investigate the benefits of such a programme over time to see when

screening benefits would become detectable to an observer monitoring the programme. This is an important question because in the current climate of government prioritising of health spending, such monitoring would be a key part of ensuring continued public funding. Because the benefits of screening are not seen until some time after the commencement of screening, it is important to know when they are likely to be first detectable and when the program should reach its full potential. The published screening randomised controlled trials showed mortality benefits after 8 to 10 years of screening. However, a programme applied to the whole population may be able to detect such benefits sooner because of the larger number of people involved.

2. How should such screening be carried out?

Colorectal cancer screening can be done using a number of different screening modalities. My study will examine the cost-effectiveness of the different modalities. I will also examine screening at different ages and using different screening intervals to arrive at the best combination of these for each modality.

3. What more do we need to know to design cost-effective screening programmes?

The aim of this modelling approach is to bring together current knowledge of colorectal cancer and screening modalities to aid programme design. However, our current understanding of cancer and screening is incomplete and one outcome of the systematic approach engendered by modelling is to highlight those areas where current knowledge is weakest. Hence one of my aims will be to identify areas where further study could improve our ability to design effective screening programmes.

1.2 Background

Screening has long been regarded as a cost-effective and clinically useful approach for early diagnosis of some cancers such as, for example, cancers of the breast and cervix. There are already organised Australian population screening programmes for breast and cervical cancer screening and some authorities have advocated such a programme for colorectal cancer (Macrae, 1996).

Population screening for cancer can consume a large proportion of the limited resources available to the health sector. Therefore it is important to assess screening outcomes to evaluate whether it is a reasonable use of these resources. However, attempting to evaluate screening procedures by the usual observational studies and controlled trials is expensive and time consuming.

One approach to investigating the effectiveness of screening is to develop a mathematical model of the disease and the screening process. This model can then be used to estimate the effect of screening on mortality and morbidity. Such models synthesise the known epidemiological data with known and projected population characteristics. They can provide timely estimates of screening effectiveness to enable health planners to make best use of scarce resources.

1.2.1 What is screening?

Screening for disease control can be defined as the examination of asymptomatic people in order to classify them as likely or unlikely to have the disease that is the object of screening. People identified by a screening test as likely to have the disease are then further investigated to arrive at a final diagnosis (Morrison 1985). The objective of screening is the early detection of a disease where early treatment is either easier or more effective than later treatment.

Figure 1.1 is a schematic representation of the main features of the natural history of a disease which are relevant to screening. The *preclinical* phase of the disease is the phase in which a person has the disease but does not have any clinical symptoms and is not yet aware of having it. Screening aims to detect the disease during this phase. In principle, the preclinical phase starts with the beginning of the disease, but in practice modelling focuses on the phase commencing at the earliest point at which the disease is detectable with a screening test. This is known as the *detectable preclinical phase*.

The preclinical phase finishes with the *clinical surfacing* of the disease. This is the point where the person develops clinical symptoms of the disease, seeks medical attention for these symptoms and the disease is diagnosed. The disease then enters the *clinical phase*, where the person has a diagnosable case of the disease.

Each of the preclinical and clinical phases may be divided into stages. For cancers these generally correspond with prognostic tumour staging schemes. The time spent in either phase or in one of the stages is designated the *sojourn time* in that phase or stage.

The outcome of a screening test is designated either *positive* if the person is identified as likely to have the disease or *negative* if they are not. All screening tests are open to error either from the test itself or its interpretation. These errors are designated as *false positive*, where a person without the disease has a positive screening result, and *false negative*, where a person with the disease has a negative screening result. The *sensitivity* of a screening test is the probability that a person with the disease has a positive screening result. The *specificity* of a screening test is the probability that a person without the disease has a negative screening result. Cases of the disease which clinically surface following a false negative result (ie. where the screening test missed the disease) are known as *interval cases*.

It is important to note that sensitivity and specificity are not properties of the test alone. For example, mammography is used to screen for breast cancer in women. In this case the sensitivity and specificity will depend on characteristics of the test, such as the nature of the mammography machine and the number of views taken, as well as on factors such as the skill of the person interpreting the mammogram, the size of any tumour in the women being screened, the density of her breast tissue and so on.

The *reliability* of a test is its capacity to give the same result, either positive or negative, on repeated application in a person with a given level of the disease. The *survival time* is the length of time between disease diagnosis, either by clinical

surfacing or detection by screening, and death. The *lead time* is the time between the detection of a disease by screening and the point at which it would have clinically surfaced in the absence of screening.

The lead time is an important issue in the examination of screening benefits. The immediate focus of screening is to detect an early form of the disease. Hence some studies have used the lead time as an index of benefit in its own right (see, for example, Zelen & Feinleib, 1969). However, this must be treated with caution. Early detection of a cancer does not necessarily confer a benefit unless there is evidence linking the associated early treatment with improved prognosis.

Lead time is also important in examining survival benefits conferred by screening. A simple comparison of survival times between screened and unscreened populations is likely to show spurious screening benefits since the survival time for a screen detected disease includes the lead time while that for a disease which surfaced clinically does not.

There is another, more subtle, reason why such survival comparisons may be spurious, even if adjusted for lead time. Screening will tend to detect people with a longer preclinical phase. This is known as *length-biased sampling*. Usually this will equate to a more slowly progressing disease. Since the disease behaviour before clinically surfacing is likely to be correlated with that after surfacing, this is likely to result in screen detected diseases having a longer survival time than clinically surfacing diseases.

Because of such issues as length-biased sampling, most screening is monitored by focussing on the ultimate disease outcome. That is, screening programmes are usually assessed by comparing overall disease mortality in the presence and absence of screening. This can be done either by comparing the screened group with a control group, as in a randomised controlled trial (RCT), or by comparing mortality before and after the introduction of a screening programme. The latter approach may be done by projecting cancer mortality based on trends prior to the introduction of the programme and comparing these projections with the observed mortality in the presence of screening.

1.2.2 Why screen for colorectal cancer?

If we exclude non-melanocytic skin cancer (which has a relatively low mortality rate and is not monitored on a routine basis by Australian cancer registries), colorectal cancer is the most frequently occurring cancer in Australia (AIHW & AACR 1999). Most colorectal cancers go through a long detectable preclinical phase as a polyp on the bowel wall (Morson 1974). If the polyp is detected and removed then the cancer is effectively prevented. Even after the cancer is established, if it is detected in an early stage it has a considerably better prognosis than if it is detected at a late stage. All this makes colorectal cancer an ideal candidate for screening.

There are a number of screening tests for colorectal cancer ranging from the cheap and moderately effective faecal occult blood test (FOBT) through to the expensive but highly effective colonoscopy. When my study commenced, in 1992, there was no

conclusive evidence that any of these screening tests would be effective in reducing overall colorectal cancer mortality in a population screening programme. Since then there have been at least four large RCT's of FOBT screening which have shown clear evidence of overall colorectal cancer mortality reduction (Mandel et al. 1993, Hardcastle et al. 1996, Kronborg et al. 1996, Kewenter et al. 1991).

Until recently there has been no general consensus on how colorectal cancer screening should be implemented but there are now a number of published screening guidelines. Winawer et al. (1997) published an authoratitive set of clinical guidelines for colorectal cancer screening on behalf of a number of American cancer and gastroenterological societies. While they make recommendations on how each screening modality should be used for both people at average risk and at increased risk, they do not address the issue of the most cost-effective combination of modalities on which to base a population screening programme. The National Health and Medical Research Council (NHMRC) published a set of guidelines for the prevention, early detection and management of colorectal cancer in 1999 (NHMRC 1999). While these also contain recommendations for colorectal cancer screening, they do not directly address the issue of a population screening programme either. However, both sets of guidelines make clear the importance of colorectal cancer screening.

1.3 Thesis structure and methods

1.3.1 Study outline

The study is based on a microsimulation model of the natural history of colorectal cancer and a model of the screening process. Chapter 2 presents a discussion of cancer modelling and a rationale for using the microsimulation approach. It also contains a discussion of my approach to model validation. Chapter 3 develops a natural history model for colorectal cancer and chapter 4 develops a screening model. Chapter 4 also presents the results of validating my model against the published RCT data.

The main focus of the modelling is FOBT screening using a guaiac test such as Hemoccult II. This is because this is the form of screening which has been most widely studied. In particular, the published RCT's provide a direct source of evidence for the benefit of FOBT screening as the primary screening modality in a population screening programme. They also provide data against which to validate the model's results. As noted in chapter 2, model validation, preferably against a RCT, is particularly important for the credibility of the model's results. Chapter 5 presents a detailed examination of a population screening programme based on FOBT screening with colonoscopy follow-up of positive results.

Chapter 5 also presents the results of simulated population screening programmes based on other forms of FOBT screening and a combination of FOBT and sigmoidoscopy screening. These have been less widely studied, so their role in a

population screening programme is less well understood. There are also no published population screening data against which to validate these models. Hence these results should be regarded as more indicative than definitive. However, they do highlight some promising alternatives to the standard FOBT screening which should be further examined.

Chapter 6 presents the results of simulated population screening using colonoscopy and sigmoidoscopy as the primary screening modalities. Again, there are no published studies of population screening based on these modalities which present data in sufficient detail for model validation. However, the clinical use of these modalities has been widely studied, so the main uncertainty here is the achievable screening participation rates.

Most discussions of colorectal cancer in the literature suggest that people at higher risk of cancer should be screened more intensively than people with no evidence of more than average risk. Both Winawer et al. and the NHMRC guidelines provide recommendations about screening people at increased risk. However, none of the major published screening studies have addressed the effect of targeting people at increased risk as part of a population screening programme. Chapter 7 simulates such a programme and examines the benefits and costs of targeting specific risk groups.

Chapter 8 draws the thesis results together and summarises the answers provided by my study to the questions posed in the study's aims.

1.3.2 Study methods

The method of microsimulation will be discussed in detail in chapter 2. All simulations in this study use 1990 as the first year of the screening programme (i.e. the *base year*). 1990 was chosen as the base year because throughout most of the study this was the most recent year for which national cancer incidence data were available. The Australian Institute of Health and Welfare has since published national data up to 1996 (AIHW & AACR 1999), but my disease incidence model is still based on national data for the period 1987 to 1990.

Population based mass screening

Each simulation is applied to a synthetic population with the age-sex structure of the whole Australian population in the base year. As each age-sex cohort is aged, they are offered screening when they reach the youngest screening age and cease screening when they reached the oldest, if an oldest screening age is specified. Colorectal cancer deaths are counted for the entire population irrespective of whether they were currently within the screening age range.

The screening programme is assumed to start in the base year and run for 40 years. There is no uniquely optimal length of time to follow a screening programme in a simulation. If the period is too short, then the simulation misses the delayed benefits of screening, such as future deaths averted by the removal of a polyp. If the period

is too long, then the simulation may become unrealistic. For example, it seems unlikely that screening and treatment protocols will remain unchanged indefinitely.

Future costs and benefits are reduced by discounting back to the base year, so that costs accrued at the end of a long programme are less relevant than those accrued at its beginning. This has the advantage that the time lag between when costs accrue and when benefits arise has less effect on a cost-effectiveness ratio calculated over a longer screening period. This is because the costs at the end of the programme, which relate to benefits occurring after the program, are reduced by discounting and so have less effect on the cost-effectiveness ratio.

My choice of a 40 year simulation period allows the youngest cohort in each of the protocols to reach the oldest age at which screening is offered in those FOBT screening protocols with a limited upper age and hence to experience all of the programme's benefits.

The costs are calculated over the screening period and discounted back to the base year. Both average and marginal costs are calculated for each screening protocol. The average cost is calculated by summing the total costs for a simulated application of a screening protocol and dividing this by the simulated years of life saved. The marginal cost of moving from one protocol to another (for example, screening people age 45 to 49 in addition to screening people aged 50 to 74) is calculated by dividing the difference in costs between the two protocols by the difference in years of life saved. Hence, in the example, the marginal cost would represent the extra cost of each extra year of life saved by increasing the screening age range (see, for example, Drummond et al. 1987).

The benefits are calculated by counting the deaths due to colorectal cancer in each of the 40 years with and without screening and calculating the years of life lost due to each death. The benefit measures are taken as the number of deaths prevented by screening and the number of years of life saved by screening (YLS), or equivalently the change in years of life lost due to colorectal cancer (YLL). The YLS are discounted back to the base year, but the count of deaths is left undiscounted so that it can be used to estimate reductions in the risk of colorectal cancer death. The results are presented in proportional terms—i.e YLS are presented as a proportion of YLL and colorectal cancer deaths prevented are presented as a proportion of total colorectal cancer deaths.

In the simulation of the RCT's used for model validation I also simulate a 95% confidence interval around the point estimate of screening benefit. It is impractical to run sufficient simulations of the whole Australian population to generate such a confidence interval for a population screening programme. Instead, ten simulations of the base model are done and the ranges of the estimates of YLS and deaths prevented calculated. This range gives an approximate idea of the variability in the simulation results due to the stochastic nature of the microsimulation approach.

The benefits of screening for screening participants

The previous section focuses on costs and benefits of population screening in terms relevant to health policy making. Its results can be used to address issues such as

whether or not support for screening for colorectal cancer is a sensible use of the resources available for health compared to other population based health interventions. However, these projected benefits relate to the population as a whole and hence underestimate the benefits which would accrue to a person taking part in such a screening programme. This section addresses the issue of the benefits of screening to screening participants. The focus of this analysis is the reduction in risk of death from colorectal cancer due to participation in a screening programme.

The average benefits of screening to screening participants are calculated by simulating screening for a synthetic cohort of one million men and one million women who are at the youngest screening target age and at age 70 at the start of the screening programme with follow-up over the rest of each population member's life. The counts of deaths and YLL are weighted by the proportion of males and females in the Australian population at the selected ages and summed to give overall values. The screening participation rate is set at one, so that all members of the population participated in all the offered screening tests. The aim is to examine benefits both to those who participate in the whole screening programme and those who, by commencing screening at a later age, only participate in part of it.

The cohort size was chosen to be sufficiently large that chance variation in the simulation results could be neglected. As an illustration, I estimate the percentage fall due to screening in cumulative colorectal cancer mortality and YLL for ten simulations of synthetic cohort at the younger age at the start of screening. Again the range of these estimates gives an approximate idea of the variability in the simulation results due to the stochastic nature of the microsimulation approach.

1.3.3 The benefits of screening over time

An important issue for monitoring the results of a screening programme is how long it will take from the introduction of screening till the benefits of screening become apparent. The benefits from screening largely arise from a shift to earlier detection, so a monitoring system could in principle assess the screening benefits by monitoring this shift. However, there are no national Australian cancer data collections which include stage at diagnosis, so any such monitoring would require the establishment of a new national data set. Further, the ultimate benefit of screening is a decrease in cancer deaths and this can be monitored directly from existing national data collections. Hence such national monitoring is most likely to concentrate on detecting and monitoring a fall in colorectal cancer mortality. This fall in mortality is the focus of the analysis in this section.

I simulate the fall in both annual and cumulative mortality in the years following the beginning of a population screening programme. These are then both plotted for the period of 40 years from the introduction of screening.

In practice a national monitoring system would only observe deaths in the presence of screening, so any gains due to screening would have to be estimated. Poisson regression has been advocated for monitoring trends in mortality rates (see, for example, Brillinger, 1986) and has been adopted for mortality surveillance by the

Australian Institute of Health and Welfare (Bennett et al. 1992). Hence poisson regression would most likely be used for national monitoring of colorectal cancer mortality in the presence of a screening program. The colorectal cancer mortality rate without screening can be projected for comparison with the observed rate with screening by fitting a poisson regression model to the mortality rates in the years prior to screening.

Bennett et al. suggest fitting a poisson regression model to 12 years of data, so I fit such a model to the age-sex standardised colorectal cancer mortality rates from 1978 to 1989 and project this forward to 1997. I then apply the simulated annual mortality falls from my model to the actual mortality rates for the period 1990 to 1997 and compared these with a one sided 95% projection interval from the poisson regression. The point where the simulated mortality with screening falls below the 95% projection interval is taken as the point at which screening benefits first become detectable.

1.3.4 Suitability of screening for government funding

The analyses presented here raise the issue of whether or not a screening programme would be likely to attract government funding. There are two approaches to this question—a comparison of the costs and benefits of screening with other health interventions which have been funded and an examination of the implicit value of health gains suggested by other government funding decisions.

George et al. examined the consistency of recommendations to list a drug on the Pharmaceutical Benefits Schedule (PBS) with the evidence on cost-effectiveness submitted by drug companies to the Pharmaceutical Benefits Advisory Committee (PBAC) (George et al. 1998). They concluded that decisions made on the basis of incremental cost per life year gained seemed to be consistent with a government willingness to pay at least about $37,000 per life year gained, at 1995–96 prices. However they found very few approvals at more than about $69,000 per life year gained. For drugs which fall between these values, the demonstrated cost-effectiveness did not appear to favour either listing or rejection. Instead the decision as to whether or not to list a drug was probably influenced by other factors. They also concluded that these data implied a value of a whole life of not less than $300,000 and up to $5.5 million.

If I adjust these to 1989–90 prices using the total health price index published by the Australian Institute of Health and Welfare, I get a range of around $32,000 to $59,700 per life year gained and $260,000 to $4.8 million per death averted (AIHW 1997). These data are based on a discount rate of 5% per year, while my costs per life year use a rate of 3%. However, I can use these values as a rough guide to the suitability of a FOBT mass screening programme for government funding. I will take the lower figure of the ranges presented by George et al. as being the point below which funding can be justified on cost-effectiveness grounds.

1.3.5 The microsimulation computer model

In 1992, when this study started, there were no widely available computer software packages for microsimulation. This meant that computer software to implement the model had to be developed along with the model itself.

The model was implemented as a suite of computer programs written in the C language. Appendix A of this thesis is a detailed discussion of these computer programs, along with a rationale for choosing the C language.

The microsimulation model depends heavily on having a good pseudo random number generator. Appendix B of this thesis is a description of the generator which was chosen for my model, along with the reasons for its choice.

1.4 Figure

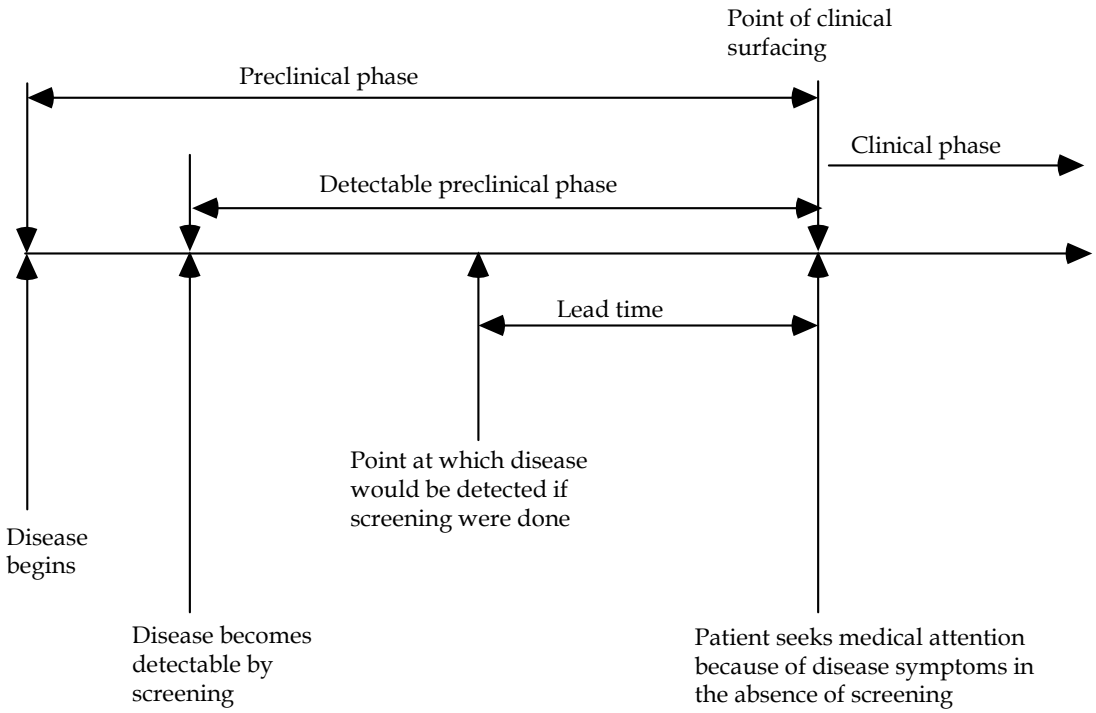

Figure 1.1: The natural history of a disease with and without screening

2 Models for cancer screening

2.1 Introduction

Mathematical modelling is a useful tool in planning and evaluating screening programmes. The purpose of this chapter is to describe the nature of the models used for cancer screening and how they have developed in recent years and to provide a rationale for the use of microsimulation modelling in this thesis.

Extensive work has been done on quantitative theories of carcinogenesis which deal with subjects such as modelling tumour growth and response to chemotherapy. These models have little application to cancer at a population level and in particular to screening programmes and hence are beyond the scope of this study. A discussion of these models can be found in Thompson & Brown (1987).

There is also a rich literature describing the use of biological models to analyse epidemiological data. These are models which describe carcinogenesis at the level of individual cells but which can be used to draw population level inferences. The most widely used of these are based on Armitage and Doll's multistage model of carcinogenesis (Armitage & Doll 1961). However, these models are not appropriate for analysis of cancer screening programmes and hence are also beyond the scope of this study. Thomas (1992) presents a general discussion of these models and Day (1990) provides a specific discussion of the Armitage-Doll model.

This is not intended to be an exhaustive study of all modelling of cancer screening. Rather it is intended to be a description of the main approaches used in modelling and their strengths and weaknesses. Shwartz & Plough (1984) present a detailed review of cancer screening modelling, while Prorok (1986) discusses modelling for cervical cancer and Prorok (1988) discusses modelling for breast cancer.

2.2 Why use modelling?

The evaluation of screening usually focuses on whether or not the screening programme has led to a fall in mortality from the disease in question. As with most medical interventions, randomised controlled trials (RCT) provide the most satisfactory empirical basis for evaluating screening programs. However, they do have significant limitations.

RCTs for screening are expensive and time consuming to run—typically requiring very large sample sizes and having long time lags till benefits are apparent. For example, the RCT of mammography screening carried out in the two Swedish counties of Kopparberg and Ostergotland had a total sample size of 134,867. A statistically significant mortality differential between the control and study groups did not appear until after six years of follow up, with a further four years of follow

up before the results could be considered definitive (Tabar et al. 1989). Twenty years of data would be required to yield results on some aspects of screening programme design (Day & Duffy 1996).

Any one trial cannot address all the issues involved in designing a screening program. For example, the Minnesota Colon Cancer Control Study used an RCT to demonstrate a statistically significant fall in mortality due to screening with a Faecal Occult Blood Test (FOBT) followed by colonoscopy in those with a positive screen (Mandel et al. 1993). However, Lang and Ransohoff have subsequently suggested that the sensitivity of FOBT is considerably less than that reported in the Minnesota study. FOBT has a high false positive rate and they argue that one third to one half of the fall in mortality could be due to chance selection for colonoscopy where an early cancer or large adenomatous polyp is present but not bleeding and the FOBT is positive for other reasons (Lang & Ransohoff 1994). The original trial's authors dispute some of the key assumptions underlying Lang and Ransohoff's analysis. They argue that the trial's results, along with other key studies of FOBT screening, are inconsistent with a large contribution of 'random' colonoscopies to the fall in mortality (Mandel et al 1994). However, since the original trial did not provide data on FOBT alone, it provides no definitive basis for deciding on the role of FOBT separately from that of colonoscopy.

Models are one way the information on the disease and screening tests from a number of different sources, including RCTs and other clinical and epidemiological research, can be combined with known and hypothesised features of the specific population to be screened. They can be used to investigate the effect of different screening regimes on different subgroups of the population, both on disease mortality and programme costs. For example, one use of modelling has been to investigate the inclusion of different age groups in the population to be screened. They can also be used to project the future course of the disease and screening programme to evaluate the changes in costs and benefits over time.

The modelling approach does have limitations. The extra information is obtained from models only by imposing assumptions about the screening process. These include assumptions about the natural history of the disease, about the characteristics of the screening test and about the behaviour of the population under study. These assumptions can only rarely be verified, though they can be evaluated as part of the modelling process.

A further complication in making these assumptions is that the natural history of cancer is not completely understood, particularly in the asymptomatic phase which is the main focus of screening. This means that any hypothesised form of the disease model may be plausible given current knowledge but may ultimately be misleading.

2.3 Types of model

Bross et al. (1968) proposed a classification of models used to analyse screening strategies into two types: *surface models* and *deep models*. Surface models comprise

the usual statistical approach to analysis. They consider only those events that can be directly observed such as disease incidence, prevalence and mortality. Deep models, on the other hand, incorporate hypotheses about the disease process that generates the observed events. Their intent is to use the surface events as a basis for understanding the underlying disease dynamics. In cancer screening, this implies models which explicitly describe the disease natural history underlying the cancer incidence and mortality.

This modelling permits generalisation from the particular set of circumstances that generated the surface events. As a result, whereas surface models provide a basis for interpreting the observable effects of screening, deep models provide an explicit basis for determining the outcomes of screening scenarios that have not been directly studied in clinical trials (Shwartz & Plough 1984). I will focus on the application of deep models to cancer and population screening.

These models can be further grouped into two broad categories—those that describe the system dynamics mathematically and those which entail computer simulation. The first of these, designated *analytic models*, uses a model of the disease to derive direct estimates of characteristics of the screening procedure and its consequent benefits. The second, designated *simulation models*, uses the disease model to simulate the course of the disease in a hypothetical population with and without screening and derives measures of the benefit of screening from the simulation outcomes.

2.4 Markov framework for modelling

Most of the cancer screening models use an illness-death model for the disease which is developed within the framework of a *Markov chain*. A sequence of random variables $\{X_k, k = 0,1,\cdots\}$ is called a Markov chain if, for every collection of integers $k_0 < k_1 < \ldots < k_n < v$

$\Pr\{X_v = i | X_{k_0}, \ldots, X_{k_n}\} = \Pr\{X_v = i | X_{k_n}\}$, for all i.

In other words, given the present state (X_{k_n}), the outcome in the future $(X_v = i)$ is not dependent on the past $(X_{k_0}, \ldots, X_{k_{n-1}})$.

The Markov chain formulation is applied to an illness-death model in the following way (Chiang 1980). The population under study is classified into n states, the first m of which are illness states and the remaining n-m of which are death states. An illness state can be broadly defined to be the absence of illness (a *healthy state*), a single specific disease or stage of disease or any combination of diseases. In modelling cancer and screening, these states typically refer to a healthy state and preclinical and clinical stages of the disease. Here a person is considered as entering the preclinical state either at the time of carcinogenesis or at the time when the disease is first detectable by a screening modality and as entering the clinical state when the disease comes to clinical attention in the absence of participation in screening.

A death state is defined by cause of death, either single or multiple. Emigration or loss to follow up may also be treated as a death state. In modelling cancer and screening, there will typically be one death state due to death from the cancer and another due to death from any other competing cause. Entry to a terminal stage of the disease is also sometimes treated as a death state.

Transition from one state to another is determined by the transition probabilities, p_{ij}, where

$$p_{ij} = \Pr\{X_{k+1} = j | X_k = i\}; i, j = 1,2,\ldots,n; k = 1,2,\ldots$$

Death states are absorbing states, since once one reaches that state, transition to any other state is impossible (i.e. $p_{ij} = 0$, for $i = m+1,\ldots,n$ and for all $j \neq i$). The disease model is said to be progressive if, once one enters the first stage of the disease, in the absence of interventions (such as screening) and competing risks, the only valid transitions are through the remaining disease stages (i.e. $p_{ij} = 0$ for $j < i$). Because the disease is modelled using a Markov chain, the future path of an individual through the illness and death states depends only on his/her current state and the future distribution of individuals between illness and death states depends only on the present distribution and not on any past distributions.

This basic model can be varied in a number of ways. The Markov chain treats time as increasing in discrete steps corresponding to the index *k*. Thus a transition between states can only occur at discrete time intervals. Most screening models extend this to allow transitions to occur in continuous time. In this case the transition probabilities for any two points in time t_1 and t_2 are

$$p_{ij}(t_1, t_2) = \Pr\{X(t_2) = j | X(t_1) = i\}; i, j = 1,2,\ldots n$$

If $p_{ij}(t_1, t_2)$ only depends on the difference $t_2 - t_1$ but not on t_1 or t_2 separately, the model is time homogenous. The simple Markov chain described above is time homogenous. This can be varied to allow the transition probabilities to vary with time. The probabilities can also be allowed to vary with age and other relevant characteristics of the individual.

Many of the cancer screening models concentrate on modelling the sojourn times in preclinical states, since these states are the focus of most screening programs. Some of the model formulations allow the probability of transition out of a state to depend on the sojourn time.

2.5 Analytic Models

A mathematical disease model with two states was first proposed by Du Pasquier (1913), but it was Fix & Neyman (1951) who introduced the stochastic version and resolved many problems associated with the model (Chiang 1980). Their model has two illness states and two death states. The two illness states are the state of 'leading a normal life' and the state of being under treatment for cancer. The two death states are deaths from cancer and deaths from other causes or cases lost to observation.

(Chiang 1964) subsequently developed a general illness-death stochastic model which could accommodate any finite number of illness and death states.

Lincoln & Weiss (1964) were the first to propose a model of cancer as a basis for analysing serial screening, in this case screening for cervical cancer. They did not explicitly use the Markov framework described above, but their model implicitly uses a classification of the disease into two illness states—a 'healthy' state, where the disease is not detectable, and a state covering the time between when the disease is first detectable and when it is actually detected by a screening examination. They also make the simplifying assumption that symptoms never appear and all disease is detected by screening. The focus of their analysis is to derive equations to estimate the distribution of the time between the inception of the tumour and its discovery.

Zelen & Feinleib (1969) proposed a simple three state, continuous time, progressive disease model for cancer incorporating a preclinical and a clinical state. The three states are defined as:

state 1: the disease free state
state 2: the preclinical disease state; and
state 3: the clinical disease state.

A death state following state 3 is implicit in this model, but is not explicitly used because the analysis focuses on the preclinical state.

A preclinical state is defined as a state where a person has the disease but clinical symptoms have not been exhibited and the person is unaware of the disease (i.e. the detectable preclinical state). The clinical state is taken as one having clinical symptoms of the disease. The authors assume that the preclinical disease will eventually progress to clinical disease if not detected and treated. In a modification to this basic model the authors further classify the preclinical state into two, defined as

state 2a: a preclinical state where the disease never progresses to the clinical state (i.e. the sojourn time is allowed to be ∞); and

state 2b: a preclinical state where the disease is progressive and will eventually progress to the clinical state.

These are used in applying the model to cervical cancer, to allow for the possibility that some individuals with the disease in a preclinical state will never have the disease progressing to a clinical state.

Attention is confined to screening programs where the individual is screened only once. The authors use the lead time as a measure of screening benefit—i.e. the time between when the individual is detected by screening in the preclinical state and when he or she would have entered the clinical state in the absence of screening.

This approach has been generalised in a number of ways by subsequent authors, with most focussing on simple disease models and the estimation of specific screening characteristics.

Prorok (1976a & 1976b) extended the lead time estimation to multiple screens. Blumenson (1976, 1977a, 1977b) calculated the probability of terminal disease as a function of disease age and used this as a prognostic measure to evaluate screening strategies. Shwartz (1978a, 1978b) modelled disease progression for breast cancer using tumour size and number of axillary lymph nodes involved to define states 2 and 3. He then determined screening benefit measures, from data on five year survival rate and five year disease recurrence rate for patients, as a function of tumour size and lymph node involvement.

Albert and his co-workers (Albert, Gertman & Louis 1978, Albert et al. 1978, Louis et al. 1978) developed a comprehensive model for the evolution of the natural history of cancer in a population subject to screening and natural demographic forces. In it's general formulation the model uses Zelen and Feinleib's classification of the disease into preclinical and clinical stages, but classifies the preclinical stage into k states which are intended to correspond roughly with prognostic tumour staging schemes. It also has two death states which correspond to clinical surfacing of the disease or death from a competing risk.

The model is progressive, in that once the initial disease state is entered the only valid transitions are to further disease states or to a death state, but allowance is made for staying indefinitely in any given state.

This model is then applied to breast and cervical cancer. Breast cancer is modelled with two illness states, state 1 corresponding to disease with no lymphatic involvement and state 2 corresponding to disseminated disease (the contrary case). Cervical cancer is modelled with three illness states, state 1 corresponding to neoplasms in situ, state 2 corresponding to occult invasive lesions and state 3 corresponding to frankly invasive lesions.

The authors then impose on this model a screening strategy with a particular probability of a positive screen, depending on a person's age and disease state. Using this they derive explicit equations describing how the natural history of cancer (depicted by the distribution of numbers in each state and associated sojourn times) evolves over time in the presence of screening. These in turn are used to derive explicit equations for measures of benefit from screening in terms of the disease status. These benefit measures include the percentage reduction in the cumulative number of observed cases of late disease due to screening and the percentage decrease in lost *salvageables* due to screening. Here a salvageable is a person who would have benefited from screening but who, in the absence of screening, progresses to a late stage of the disease before discovery.

Dubin (1979, 1981) developed a general multi stage disease model similar to that of Chiang (1964) and applied this to breast cancer using the same two stage classification as Albert et al. (1978). He noted the difficulty in estimating parameter values for detailed disease models from existing data from screening programs. His model aimed to avoid these difficulties by maintaining comparability between the model and the observable characteristics of a screened population. He did this by focusing on age and stage specific incidence and survival times in the presence and absence of screening. He derived formulae for the proportion of disease incidence

which has been diagnosed earlier due to screening than it would have been in the absence of screening and used these formulae to derive various measures of screening benefit. Dubin's model is not strictly a deep model as defined above. However, although he makes no explicit hypotheses about the rate of disease progression, such hypotheses are implicit in his model.

Day & Walter (1984) developed a variation on the simple three stage model which has been extensively used. The focus of this model is the sojourn time in the preclinical state, for which a probability distribution is specified. For example, Walter & Day (1983) in applying the model to breast cancer, used several alternative distributions including the exponential, the Weibull and a non-parametric step function. Under the model assumptions, one may derive expressions for the anticipated incidence rates of clinical disease among groups with particular screening histories and for the anticipated prevalence of preclinical disease found at the various screening times. One advantage of this model is that it is relatively simple to obtain approximate confidence intervals for parameter values. The model was extended by Walter & Stitt (1987) to permit evaluation of survival of cancer cases detected by screening.

A useful synthesis of the analytic models described above applied to breast cancer is presented by O'Neill et al. (1995).

All of the above are progressive models but there are some forms of cancer for which the assumption of progression is not appropriate and for which some form of regression is required. These are cancers, such as colorectal cancer and cervical cancer, where screening detects preinvasive or even precancerous lesions (Brookmeyer & Day 1987).

A number of models have attempted to address this. Coppleson & Brown (1975) used data on age specific clinical incidence and detection rates of a first smear to fit a four state model for cervical cancer. They found that the observed data could not be explained without allowing for regression. Albert (1981) developed a variation of his earlier model for cervical cancer which allowed for regression from the carcinoma in situ stage back to the healthy state. Brookmeyer & Day (1987) and van Oortmarssen & Habbema (1991) both developed similar extensions to the Day and Walter model to divide the preclinical stage into two. The first stage allows regression to a healthy state but once a cancer reaches the second stage only progression is allowed.

The van Oortmarssen and Habbema model provides an interesting variation on the use of these models. In their study the aim of the model was not to study cancer screening directly. Rather, the model was used to study the disease dynamics and in particular to examine the epidemiological evidence for the existence of regression in pre-invasive cervical cancer.

The models described above follow a common theme of characterising the disease as a series of states (corresponding to health, the various disease stages and death) with people moving between the states with certain transition probabilities and/or certain sojourn times. Screening is then evaluated by superimposing on the disease

process a strategy with a screen of some sensitivity. This is in contrast to the next model, due to Eddy (1980), which uses a different strategy.

Eddy's modelling strategy uses a time varying Markov framework. However, he models the interaction between the screen and the disease in his basic model. His basic model is a four stage one defined in terms of three time points. The first is a *reference time point* t_p. The way this is defined varies with the cancer under discussion but, as an example, for breast cancer it is the point at which the disease can first be detected by physical examination. The *occult interval* is then defined as the time interval between this and the point t_M at which the disease is first detectable by screening (e.g. by mammography). The *patient interval* is defined as the time between t_p and the time t_Π at which the patient would actually seek medical care for the lesion.

With Eddy's model t_Π, t_p and t_M can occur in any order. The important assumption is that once a disease is detectable by a screening modality (ie after t_M) then any screen using that modality will always detect the disease. This assumption replaces the normal assumption that successive screens are independent. The other two states are a 'healthy' state (which includes any preclinical disease which is still undetectable by screening) and a clinical disease state. Eddy models the probability distributions of the occult and patient intervals and uses these to derive formulae for the probabilities of discovering a malignant lesion by screening and by other methods.

Eddy's model has been applied to several breast cancer screening data sets as well as to cervical cancer and gastrointestinal cancer. It has also been extended to the case where there is more than one type of screening test (Eddy et al. 1987).

Finally I shall describe three recent analytic models which provide interesting variations on the basic disease model.

The first of these is the stage shift model (Connor et al. 1989). This model allows any number of illness stages and assumes that the effect of screening is to shift the diagnosis of a cancer from a higher to a lower stage or, within a given stage to an earlier time of diagnosis. The method of fitting this model requires a completed randomised controlled trial with equal sized intervention and control groups and further requires that follow up has reached the point in time where comparable sets of cancer cases have accumulated in the study and control groups. This limits its applicability, but it has been successfully used to analyse breast cancer screening data (Chu & Connor (1991).

The second is the peak analysis model (Baker et al. 1991). This uses data from a randomised trial to determine the time period when the effect of screening on mortality reduction is maximum. The results of the trial can then be analysed restricting attention to that time period, providing more powerful statistical tests. For breast cancer screening, for example, this could mean excluding the mortality experience of the first few years after the initiation of screening. A disadvantage of this model is that the selection of the peak time period for the mortality comparison could be regarded as 'data-driven' and subject to the usual problems of a post hoc analysis (Miller et al. 1990).

The third is the use of surrogate end points for RCTs to shorten the duration of the trial and to increase the power (Day & Duffy 1996). Day and Duffy apply this approach to a study comparing breast cancer screening at three yearly and one yearly intervals. Tumour size is the most important variable in predicting survival from breast cancer in the screening context, so they consider the difference in tumour size distribution between the study groups. They show that using this as an index of benefit and projecting expected mortality allows a result after only five years, compared with the 15–20 years required for a trial based on observed mortality. Further they demonstrate the rather surprising result that the use of surrogate endpoints leads to an increase in the power of the RCT compared with using the observed mortality. While completed trials remain necessary to establish the primary benefits of screening, this approach allows faster and more efficient resolution of subsidiary issues.

2.6 Simulation models

Knox 1973 developed the earliest and most comprehensive simulation model. As with the analytic models, Knox uses a healthy state, a number of illness states and two death states. However, the model involves considerably more illness states, including classifying the disease as a preclinical, early clinical or late clinical cancer, and further classifying each of these as treated or not treated and each cancer as high or low grade. The result is a model with 26 defined states, though not every state was used for every simulation.

Knox defines a *transition matrix* which gives the estimated transfer rates between the various pathological states, modified suitably according to the age of the individual or the duration of the state. He then simulates the evolution of the disease in a hypothetical cohort of study subjects which has similar characteristics to the population he wishes to study (which, in this case, is the adult female population of England and Wales) using the transition matrix and a standard life table to provide the risks of competing causes of death.

Finally he adds details of the screening procedures to be considered, specifying the clinico-pathological states to which they apply, and their sensitivities and specificities in relation to each, and the transfers between model states which will occur following detection or non-detection. The screening policies are arranged in incremental series and the results compared with each other and with the results of providing no screening at all. This allows the appraisal of benefits and costs in both absolute and marginal terms.

This model has been applied to both cervical cancer (Knox 1973) and breast cancer (Knox 1975). It illustrates one major difference between the analytic and simulation approaches. This is the greater complexity of the disease model in the simulation case. In general, simulation models are capable of considering more complex disease models and screening procedures. However, this extra complexity requires more detailed information on the disease dynamics in order to specify the model

and this information is often not readily available. Knox (1988) says of his earlier work that

'The chief problem of applying the predictions stemmed from uncertainties about the clinical course of the early stages of cancer…'

and in this and all his subsequent analyses he simplified his model to one with only two illness states. This model is worth discussing in detail, because it provides an interesting variation on the usual modelling approach.

The major difference in this approach is in the population to which the model is applied. Whereas the usual approach is to consider all people at risk of a cancer and to use the model to project mortality with and without screening, Knox's approach is to consider only those who have died from cancer. He uses the model to estimate how many would have been saved if screening had been offered. He refers to this as "tearing down" a graph of age-distributed deaths in successive steps through the insertion of screening procedures at selected ages. (Knox & Woodman 1988)

This means that Knox does not need to consider variations in the course of the disease such as lesions which never clinically surface or which regress to a healthy state, because all his population have, by definition, a progressive form of the disease.

The two illness states are designated A and B. During state A the disease is susceptible to early detection and full or partial cure. During state B the disease is incurable. The sojourn time in each state varies around an age specific mean. The screening procedure has a probability of detecting the lesion which rises linearly during period A, while the probability of curing the disease falls linearly during A.

This model has the advantage of simplicity, which means that it is relatively easy to find plausible parameter values for it. However, this simplicity has disadvantages. The model only considers the situation of a fully established screening program, so that it cannot be used to investigate issues surrounding setting up a new program. Also, because it is focussed on mortality reduction, it cannot be used to consider issues relating to costs of screening programs.

Researchers at the Australian Institute of Health and Welfare have extended this approach by combining Knox's disease model with a costs model to evaluate the introduction of breast and cervical cancer screening programs in Australia (AHMAC 1990, AHMAC 1991). They have also combined the disease model with mortality projections to investigate the timing of mortality reductions due to the introduction of a breast cancer screening programme (AIHW 1992, p83).

Parkin (1985) identifies a number of advantages of the usual simulation approach of transferring year by year specified proportions of a single cohort in a deterministic fashion between model states. These include the model's ability to

- demonstrate the relationships between variables;
- explore the effects of different acceptance rates and test characteristics on outcome measures;

- examine the net cost-effectiveness of different screening policies by imputing costs to the different outcomes of screening tests; and
- explore the effect of different theoretical natural histories on the outcome of screening.

However, he also identifies some of the disadvantages of this approach. Firstly, in practice, services have to be planned not for a single cohort over an entire life span, but for a very heterogenous population over relatively short time periods. When a screening programme providing for testing at certain fixed ages is introduced into a community, only people younger than the starting age for the screening policy can possibly receive the full schedule of tests. Thus, benefits from screening will at first be small, but increase progressively as more of the population receives a series of examinations. Further, many people will have already had previous examinations so the results of the screening policy will depend on the existing screening status of the population. This cannot be simulated by a single cohort model, nor can differences in the risk of disease in different birth cohorts.

Secondly, it may be desirable to use characteristics other than age to identify subgroups of the population for selective screening. This is less often of practical use, since such subgroups are often not readily identifiable. However, a planning model should be able to explore the effectiveness of policies involving differential screening of such sub populations. In addition, population subgroups often have different rates of attendance at screening programs which may be correlated with different disease risks.

Finally, in real life, screening programmes do not exist in isolation from the rest of the health care system. Much screening activity can take place outside a screening programme. Most models usually treat this activity as "diagnostic" and ignore it. However, a planning model should take account of all relevant screening activity.

Parkin proposes instead a *microsimulation* approach. Here the life histories of individual members of a population are simulated. The population in his model has the demographic makeup of that of England and Wales and it's size is governed by two considerations: (i) the computer time involved in microsimulation of very large populations and (ii) the need for reliable results in a stochastic simulation of relatively rare events.

Each individual is characterised by his or her values for a set of variables which will be used in simulating demographic events, disease natural history or screening programmes. The values of these variables are updated annually using sets of conditional transition probabilities (e.g. the probability of childbirth given age, marital status and initial parity). The occurrence of a transition is decided by comparing the relevant probability against a randomly generated number.

The disease model used is of the standard form. The cancer is conceptualised as having a number of states, including a health state and a death from other causes state.

There is considerable flexibility in modelling screening programmes and, since the model follows individuals, it is possible to simulate contacts with the health care system and the "incidental" screening which occurs on such occasions.

Parkin's microsimulation model was developed specifically for cervical cancer screening, but a group working at Erasmus University in the Netherlands has developed a general modelling framework for microsimulation modelling of cancer screening called MISCAN (MIcrosimulation SCreening ANalysis) (Habbema et al. 1983, van Oortmarssen et al. 1990). Strictly speaking, MISCAN is not itself a model, but rather a model generator—a package which can generate and calculate a variety of these microsimulation models.

The MISCAN approach, like Parkin's model, is based on the actual structure of a population as it develops in a given country at a particular time . The mass screening programme under consideration is taken as starting in a particular year and finishing in a particular year. Standard demographic techniques are used to project the study population to a year well after the nominated end of the programme. This allows for both the introduction of the programme to be modelled and the effects after the end of the programme to be followed up.

The basic structure of the cancer model is similar to Knox's earlier model with a detailed classification of clinical and preclinical cancer states, though it uses a smaller number of states. The interaction between the disease model and the screening programme is designed to allow projection of screening and treatment costs as well as cancer mortality and morbidity.

MISCAN has been widely used to analyse breast and cervical cancer screening programmes.

2.7 Model fitting and validation

Eddy (1985) proposed four levels of validation for mathematical models. These are:

- First order validation: which requires that the structure of the model make sense to people who have a good knowledge of the problem;
- Second order validation: which involves comparing estimates made by the model with the data which was used to fit the model;
- Third order validation: which involves comparing the predictions of the model with data which was available when the model was fitted but was not used in the estimation of model parameters;
- Fourth order validation: which involves comparing the outcomes of the model with observed data when applied to data generated and collected after the model was built (for example, data from a previously unobserved screening programme).

In this section I shall discuss model fitting and validation in the framework of these levels.

First and second order validation are generally not difficult to accomplish, but their implications for model validity are difficult to interpret. The conceptualisation of cancer as a series of preclinical and clinical stages is virtually universally accepted as a reasonable characterisation of the disease. Problems may arise with first order validation when the details of the disease stages are specified but generally a wide variety of model formulations are plausible within the constraints of the very limited knowledge of preclinical cancer.

Second order validation highlights one of the central problems with this sort of deep model. This is the difficulty of directly relating available data to model parameters. The mismatch between the data available, either from screening trials or other sources, and the model data requirements for parameter estimation has been recognised from the beginning of this type of modelling. Lincoln & Weiss (1964) note, for example, that

"Here we can do no more than introduce plausible forms for the different functions involved and plausible values for the parameters" (p188)

They go on to describe the difficulties in relating available data to the mathematical functions on which their model is based. This is a recurring problem in modelling cancer for screening and to some extent affects all the models described in this study.

Some of the analytic models have developed methods of estimating model parameters using standard statistical estimation approaches. Dubin (1981), for example, structured his model so that it could directly use the data from screening trials, though as a consequence his model relates less to the disease natural history than the others. Louis et al. (1978) derived non parametric models for the probability distributions specified in their model and proposed the use of maximum likelihood methods to fit them. Day & Walter (1984) used both parametric and non-parametric functions for their preclinical sojourn time and suggest either maximum likelihood methods or least-squares criteria to fit them. However, many of the analytic models and all of the simulation models proceed in a more ad-hoc fashion by varying their disease natural history and model parameters until their models closely reproduce existing data.

An example of how this ad-hoc fitting operates is given by Knox (1975), where his earlier model is fitted to breast cancer screening data. Knox describes fitting the natural history data thus:

"A statement of the natural history of the disease process must be provided in the form of a 'transition matrix' which gives estimated transfer rates between the various pathological states, modified suitably according to the age of the woman or the duration of the state. This set of values is adjusted iteratively until an output is produced which matches available data on incidence, prevalence and mortality. If, as sometimes happens, more than one natural history statement is capable of mimicking these facts, then the natural history will have to be treated as one of the uncertainties. Subsequent runs will then have to be repeated for a range of natural history alternatives, and each prediction of results will be conditional upon the accuracy of the natural history used" (pp17–18)

Parkin (1985) provides an example of just such an uncertainty about natural history, with the final model including three different natural histories as alternatives.

This approach to model fitting means that, particularly for models with a large number of unknown parameters, the fit of the predicted values is likely to be close to the observed data whether or not the model is in any sense valid. As Eddy (1985) notes

"Since the structure and parameters of the models are estimated to predict the observations, it should be no surprise when they do" (p149)

Third order validation is usually made difficult by the lack of data. Generally most available data are used in determining the parameters of the model (Shwartz & Plough 1984). This is particularly true for breast cancer models.

The only real data source for fitting models for breast cancer screening are the screening studies and in particular the randomised controlled trials. The first major study was the Health Insurance Plan of New York study (HIP) (Shapiro et al. 1982) This programme started screening in 1963. Subsequent studies were not started for another 10 years, with the Utrecht Screening Programme (Collette et al. 1984) starting screening in 1974 and the Swedish Two-county Randomised Trial starting in 1977 (Tabar et al. 1989). This means that many of the models only had access to the HIP data. Screening technology has changed significantly since the HIP programme began (van Oortmarssen et al. 1990), so when later studies became available, they could not be directly compared with the HIP programme. In any case it is questionable whether models based only on HIP data are directly relevant to modern screening (Day & Miller 1988, p111). Because of the long time before mortality benefits from screening are fully apparent, models fitted using solely data from later studies have only appeared relatively recently (van Oortmarssen et al. 1990) and, at least in their published form, have generally not addressed the issue of third order validation.

Eddy (1985) recognised that fourth order validation is only possible in rare cases (p150). However, there are at least two examples of studies which use models in a way that could be called fourth order validation, coincidentally both using Eddy's own model. Verbeek et al. (1988) compare predictions from Eddy's model for breast cancer to data from a mammography screening programme in Nijmegen. The authors note that the comparison does not suggest too good a fit. However, this is only a preliminary study and further validation work remains to be done. Eddy (1987) compares his model for cervical cancer with a later independent analysis of empirical data. In this case the model appears to predict accurately the effect of different cervical cancer screening policies on outcomes that are important for policy decisions.

I will apply model validation to my simulation model using this framework.

2.8 Current state and future directions

The problem of model validation and its effect on the credibility of model based results is still a barrier to their wider use. Nevertheless there are a number of areas where modelling can make a uniquely important contribution to our current understanding of screening.

In the absence of specific randomised controlled trials, modelling remains the only effective way of evaluating different screening regimes. For example, at the time mammography screening was introduced to Australia, the inclusion of women aged between 40 and 50 in the programme was still a contentious issue with no international consensus on the effectiveness of screening for these women (AHMAC 1990, p70). While it could be argued that decisions on screening these women shouldn't be made in the absence of reliable evidence on the presence or absence of the benefits, in practice governments did develop screening programmes and modelling played an important role in guiding policy makers.

Modelling also has a crucial role to play in assessing the cost effectiveness of screening programmes. Even for cheap and easily available screening technologies, organised mass screening programs are the best way to ensure that the benefits of screening are fully realised (AHMAC 1991). Modelling is not only necessary in order to plan these programs, but funding bodies are unlikely to fund such programmes without at least initial cost-effectiveness studies and modelling is the only practical way to derive the necessary estimates of future benefits and costs.

Miller et al. (1990) best summarise the current situation when, in discussing some recent models, they say

"It is clear that these, and other models already developed or under consideration, may enhance our understanding of the natural history of screen-detected lesions and the process of screening. However, they require validation with the best available data, which is preferably derived from randomised trials, before they could be extrapolated in ways that might guide policy decisions. As such data become available, assumption-based models need to be modified to incorporate this extra information, in order to improve the extrapolations needed to make policy." (p 768)

While analytic models have a role in investigating specific facets of the disease and screening process (see for example van Oortmarssen & Habbema, 1991), the more comprehensive simulation models, and particularly the microsimulation models, seem best suited to the overall assessment of costs and effectiveness in screening programmes and the investigation of different screening regimes. However, the challenge in using the simulation approach is to derive disease and screening models which are sufficiently complex to model all relevant aspects of screening but sufficiently simple to enable interpretable second order validation.

3 A model for the natural history of colorectal cancer

3.1 Natural history and epidemiology of colorectal cancer

3.1.1 Incidence and prevalence

Table 3.1 lists the most frequently occurring cancers in Australia in 1996, the most recent year for which national cancer incidence data are available. If we exclude non-melanocytic skin cancer (which has a relatively low mortality rate and is not monitored on a routine basis by Australian cancer registries), colorectal cancer is the most frequently occurring cancer in Australia. There were 10,998 new cases of colorectal cancer in 1996, which is 14.2% of all new cancer cases. Colorectal cancer is the second highest cause of cancer death after lung cancer—causing 4,606 deaths in 1996, which is 13.5% of all cancer deaths (AIHW & AACR 1999).

As shown in figure 3.1, the incidence of colorectal cancer is low among people aged under 50 but rises steadily after this age. It occurs less often in women than in men—in my base year men had an estimated lifetime risk of 1 in 19 of developing the disease compared to a lifetime risk of 1 in 27 for women (Jelfs et al. 1996).

Figures 3.2 and 3.3 present recent trends in colorectal cancer incidence and mortality. Incidence has risen slightly for men but fallen for women. Mortality has fallen for both men and women. The reasons for these changes are not clear, but the decreases may be due to a combination of better diet and some use of screening.

Colorectal cancer can be classified according to its primary site as being in the proximal colon, distal colon or rectum. Table 3.2 presents this classification in terms of the ICD-9 coding of the disease. Figure 3.4 presents the proportion of colorectal cancers with a primary site in the proximal colon. This shows an increase in this proportion with age after the age of 40, a result which is in line with previous international studies (Cooper et al. 1995). The reasons for this are not clear, but Cooper et al. argue that the most likely explanation is biologic differences between the proximal and distal colon which effect carcinogenesis (Cooper et al. 1995). The main implication for my study is that screening modalities which do not reach the proximal colon (such as sigmoidoscopy) may be less effective at older ages.

3.1.2 Clinical features

Symptoms and signs of colorectal cancer may include abdominal pain, change in bowel habits, bleeding, an abdominal or rectal mass, abdominal tenderness and

evidence of bowel obstruction (Winawer et al. 1997). The problem in clinical diagnosis is that the symptom patterns are very variable and are also common in people with other diseases such as, for example, haemorrhoids or fissures and irritable or spastic colon (Clark et al. 1987).

Colorectal cancer is classified into stages according to the extent to which it has extended from its origin in the mucosa through the wall of the bowel, to regional lymph nodes and to distant sites (especially the liver) (Winawer et al. 1997). I will follow the majority of studies of this cancer in using Dukes' staging system. Table 3.3 presents the Dukes' stages compared with the TNM system. The key features of Dukes' stages are:

Stage A: The lesion is confined to the mucosa (the surface layer of cells);

Stage B: The lesion has extended beyond the surface to invade the underlying muscle layer of the intestinal wall (*muscularis propria*) but there are no signs of lymph node involvement;

Stage C: There has been some invasion of the lymph nodes;

Stage D: There are metastases to other organs;

Survival is closely related to the cancer stage at diagnosis (Newland et al. 1981). Generally tumours in stages A and B have a relatively good prognosis, those in stage C have a poor prognosis and those in stage D are regarded as incurable. For example, the five year survival rate for colorectal cancers recorded on the South Australian Hospital Registries varied from around 90% for stage A to lower than 10% for stage D (South Australian Health Commission & South Australian Cancer Registry 1994). These South Australian data also show that tumour differentiation and, for rectal cancers, age at diagnosis and residence outside Adelaide influence survival. However, the dominating factor is stage at diagnosis. The dominant place of stage at diagnosis in prognosis is also supported by international research (Stryker et al. 1987).

3.1.3 Risk factors

Dietary factors and physical activity have been identified as major risk factors for colorectal cancer, so that an estimated 66% to 75% of new cancer cases could be prevented by diet and physical activity (NHMRC 1999). However, because of the complexity of the relationships between these risk factors and the disease, it would be impractical to involve the risk factors when defining a target population for screening. Hence the discussion presented here focuses on the risk associated with disease family history and specific genetic syndromes.

About 75% of all new cases of colorectal cancer occur in people with no known predisposing factors for the disease (Burt et al. 1990). These people are considered to be at average risk of the disease.

The remaining cases occur in people who are at higher than average risk of the disease. This group comprises people with a family history of colorectal cancer, people with specific genetic syndromes, people with a disease such as inflammatory

bowel disease which predisposes to colorectal cancer and people with previously diagnosed cancer. Risk factors for colorectal cancer are summarised in table 3.4. The discussion of high risk groups presented here is taken from Winawer et al. (1997).

People with a family history of colorectal cancer but without any apparent defined genetic syndrome account for most of those at high risk—about 15% to 20% of all cancer cases. Hereditary nonpolyposis colon cancer accounts for 4% to 7% of all cases and familial adenomatous polyposis about 1%. The remainder of cases among those at high-risk, about 1%, are attributable to a variety of uncommon conditions such as inflammatory bowel disease, Peutz-Jeghers syndrome and familial juvenile polyposis.

Family history

People with one or more first-degree relatives (parent, sibling or child) with colorectal cancer but without one of the specific genetic syndromes have approximately twice the risk of developing colorectal cancer as people at average risk. In those with a single affected first-degree relative, the risk significantly increases in the fourth decade and continues to increase with age. The risk is further increased if more than one first-degree relative is affected and if the relative's cancer occurred before the age of about 55. A similar increased risk is found in close relatives of people having an adenomatous polyp diagnosed below age 60. Most of this risk seems to be inherited rather than environmental.

Specific genetic syndromes

Familial adenomatous polyposis (FAP)

People affected with FAP develop adenomatous polyps in the second and third decades of life, have hundreds to thousands of polyps throughout the colon by their 30's and have an almost 100% chance of developing colorectal cancer by their 40's. Syndromes once thought to be distinct but now known to be variants of FAP include Turcot's syndrome (familial colorectal and brain cancer) and Gardiner's syndrome (familial colorectal cancer, osteomas, and benign soft tissue tumours).

Hereditary nonpolyposis colon cancer (HNPCC)

HNPCC has two main forms: one without a family history of other cancers (Lynch syndrome I) and the other with an increased familial occurrence of other types of cancers, typically of the ovary and uterus (Lynch syndrome II). In both variants of the syndrome, colorectal cancers occur at an early age, in the fourth and fifth decades. Adenomatous polyps precede the development of cancer but do not occur in unusually large numbers (as with FAP). The clinical distinction between patients with HNPCC and those with a family history of colorectal cancer in the absence of a genetic syndrome is sometimes unclear. Because HNPCC is not marked by polyposis of the entire colon or other clinically apparent abnormalities, if genetic testing is not performed the syndrome must be defined by family history. A common standard is

- the existence of three or more relatives with histologically documented colorectal cancer, one of whom is a first degree relative of the other two;
- one or more cases of colorectal cancer diagnosed before age 50 years in the family; and
- colorectal cancer involving at least two generations.

Because these criteria are strict, defining HNPCC in this way is unlikely to result in false positive diagnoses, though they may give rise to false negatives.

Hamartomatous polyposis syndromes

In some uncommon conditions there are hamartomas in the small and large bowel. The risk of colorectal cancer is increased, but the magnitude of this increase is not precisely known. In Peutz-Jeghers syndrome, hamartomas throughout the small bowel and mucocutaneous pigmentation develop in childhood and colonic adenomas may occur. In juvenile polyposis, colonic hamartomas also develop in childhood, without abnormal pigmentation and adenomatous features can occur in some of these polyps.

Inflammatory bowel disease (IBD)

Patients with IBD are at substantially greater risk of cancer than people in the general population, although this contributes less that 1% of all new cases of colorectal cancer. The risk associated with ulcerative colitis and Crohn's colitis is similar. Cancers developing in people with IBD are generally flat and infiltrating and do not often arise from polyps.

Prior colorectal cancer

People who have had colorectal cancer are at increased risk of developing another (metachronous) colorectal cancer (apart from recurrence of the original cancer). In a retrospective analysis of 5476 average-risk individuals entered in a tumour registry in Nebraska, the annual risk of a metachronous lesion was constant over 20 years from the first cancer (Cali et al. 1993). In comparison with colorectal cancer incidence data in the general population, the ratio of observed to expected cancers was 1.45 ($P < 0.01$). The risk of developing a metachronous lesion was 0.35% per year, reaching 6.3% at 18 years.

3.1.4 Polyps

Mucosal masses in the colon and rectum are described clinically as *polyps*. They represent several kinds of histology with different clinical importance. They include the very common hyperplastic polyp (10% to 30% of all polyps), which are generally small (< 5 mm in diameter) and are of no clinical importance, mucosal tags (also 10% to 30% of all polyps), which are also small and of no clinical importance, and a variety of other histological types such as lipomas and hamartomas which are uncommon (Winawer et al. 1997). The polyp which can be regarded as a

premalignant stage of colorectal cancer is the adenomatous polyp (Fenoglio & Lane 1974). It accounts for about half to two thirds of colorectal polyps.

The prevalence of adenomatous polyps will depend on external risk factors such as diet and hence is likely to vary between countries and cultures (Winawer et al. 1990). Autopsy studies in Europe and North America suggest that they are found in about a quarter of people by the age of 50 rising to over half by the age of 80—with a prevalence rate of about 60% for men and 40% for women (Winawer et al. 1997).

The *muscularis mucosa* forms the boundary between the neoplastic tissue in the adenoma and the *submucosa* and is the structure which is used to distinguish between intramucosal (in situ) and invasive carcinoma. A focus of carcinoma confined to an adenoma above the muscularis mucosa is not considered invasive cancer because it does not metastasise. The transition to invasive cancer occurs once the boundary is crossed and the carcinoma reaches the submucosa (Fenoglio & Lane 1974).

Clearly not all adenomatous polyps progress to invasive cancer. Studies of untreated polyps suggest that some polyps remain stable in size, some grow and some shrink and disappear over time. Those that either shrink or remain stable in size below 10 mm in diameter are unlikely to progress to invasive cancer. Those that grow larger than 10 mm are highly likely to progress to invasive cancer (Hoff et al. 1986, Stryker et al. 1987, Welin et al. 1963, Figiel et al. 1965, Knoernschild 1963). Autopsy studies suggest that the small and clinically unimportant polyps are uniformly distributed throughout the colon (Winawer et al. 1990).

Polyps can be classified histologically as tubular adenomas, villous adenomas or an intermediate type. The risk of invasive cancer increases with the more villous structure. The risk also increases with multiple occurrence of polyps. However, the predominant risk factor appears to be the size of the polyp (Muto et al. 1975, Fenoglio & Lane 1974).

Studies of high grade dysplasia in polyps, which is thought to be the bridge between a polyp and invasive cancer, suggest that the sex of the patient has no effect on malignant potential but age has some effect. However size and villous component are the predominant factors (O'Brien et al. 1990).

3.1.5 The adenoma-carcinoma sequence

The role of colonic polyps in colorectal cancer has been a controversial one, with some authors arguing that these polyps play little or no role in the development of cancer (e.g. Ackerman et al. 1964), while others argue that polyps form a precancerous phase for the majority of such cancers (e.g. Morson, 1974). The current consensus favours the polyp-cancer sequence for the majority of colorectal cancers (Winawer et al. 1997, Winawer et al. 1990), though the exact proportion of cancers which arise from polyps is not known.

The average time taken for a polyp to progress from small adenoma to invasive cancer and the distribution of progression rates are not precisely known, but there is indirect evidence for the likely time taken by this progression.

Morson and Muto et al. suggest that the time taken for a small polyp to progress to an invasive cancer is not less than 5 years, possibly more than 20 years but with an average between 10 and 15 years. They base this on direct observation of a small number of untreated polyps and the time lag between diagnosis of polyposis and development of cancer in patients with familial polyposis (Muto et al., 1975, Morson, 1974).

A case control study carried out by Selby et al. (1992) suggests that the protective effect of sigmoidoscopy lasts for up to 10 years. A study by Kozuka et al. (1975) showed that the average age difference between people with the earliest stage adenomatous polyps and those with invasive cancer was 18 years. While not precisely specifying the transformation time from small adenoma to invasive cancer, these studies confirm that it occurs slowly over many years.

There is also evidence that a small number of polyps with no cancerous significance also exhibit growth (Welin et al. 1963, Figiel et al. 1965).

3.2 Overview of the disease model

3.2.1 Model description

Figure 3.5 is a diagram of the disease and screening model used for colorectal cancer in this study. The disease is characterised as a series of transitions between disease states. The sojourn time in each state for each person with the disease and their transition to the next state are assumed to be independent of their path through the previous states. The state 'death from other causes' may be entered from any of the model states.

In the first state the person is free of colonic polyps and colorectal cancer. Their first transition into the disease model is assumed to be the development of a polyp. Polyps may spontaneously disappear, in which case the person re-enters the healthy state, or they may grow and lead to invasive cancer. Once the person enters an invasive cancer state the only valid transitions are to a more advanced disease state or to disease detection.

As noted above, there is good evidence that villous adenomas have a higher malignant potential than tubular adenomas. However, the size of the polyp has the paramount place in the risk of developing invasive cancer. Further, as will be discussed in the next chapter, published screening test sensitivities relate mainly to polyp size. In view of this and of the fact that there are insufficient data to model polyp growth separately for these histological types, the model presented here does not address this issue.

There are no data available on which to base an investigation of differentials in polyp growth rate by age. However, Morson's observation that patients with FAP provide a good model for studying polyp growth rates suggests that age is not a relevant factor (Morson 1974). Polyps in people with FAP occur at a much younger age than those in people at average risk. If age affected the polyp growth rate, then

FAP would not provide such a model. My model assumes that polyp growth rates are independent of age.

If the disease is detected in the polyp phase, then removal of the polyp is assumed to provide a complete cure. The model assumes that for large polyps (above 10 mm) the person is then subject to medical follow-up and any further polyps will be removed. Hence the person will remain cancer free and ultimately die from another cause.

Cure is also assumed possible from some of the invasive cancer stages. In these cases the person is also assumed to be subject to medical follow-up which will keep them free of the disease so that they ultimately die from another cause.

In both the cases of large polyp removal and cancer cure, the person is assumed to be ineligible for further participation in a population screening programme.

The model assumes that incidence and death from colorectal cancer are independent of death from other causes. This assumption may not be true. For example the risk factors for colorectal cancer include factors such as diet which are also risk factors for diseases such as heart disease and stroke. However, there are no data available to support an analysis of a relationship between this cancer and other diseases which would allow me to modify this assumption.

People with multiple cancers may have different clinical surfacing distributions and survival times to those with only one cancer. The available data do not provide an adequate basis for estimating separate model parameters for these people. Instead the model assumes that multiple cancers in a single person act independently. Each cancer is given the probability of screen detection or clinical surfacing appropriate to its stage and the probability of the person being identified as having cancer is the probability of at least one cancer surfacing or being detected. When one cancer is detected, all cancers present at that time are assumed to be detected and the person is given survival probabilities appropriate to the most advanced of the cancers.

The model presented here is developed for people who would be eligible for a population screening programme. The low incidence of colorectal cancer at younger ages means that general population screening programmes would not apply at these ages. Cancers among high risk groups defined by specific genetic syndromes occur predominantly in these younger age groups and so are beyond the scope of this model. As noted above, HNPCC is often difficult to distinguish from a family history of colorectal cancer without a specific genetic syndrome and cases of cancer associated with HNPCC may occur in the screening programme age group. There is some evidence that polyps and cancer associated with HNPCC may behave differently from other polyps and cancers (Winawer et al. 1997). However, I will assume that the numbers of such cases are small and may be ignored for the purposes of modelling the disease.

Cancers associated with inflammatory bowel disease do not follow the polyp-cancer sequence assumed by the model. However, people with these diseases are most likely to be under medical care for the disease. I will assume that they will be

monitored for the development of cancer as part of this medical care and hence are beyond the scope of my model.

3.2.2 Modelling assumptions

This section presents the main assumptions underlying the natural history model. A full discussion and justification of these assumptions is given in the relevant sections below.

General assumptions

- The probability of death from causes other than colorectal cancer applying in the base year also applies throughout the model projection period;
- The probability of death from causes other than colorectal cancer is independent of the person's colorectal cancer status.

Polyp phase

Figure 3.6 is a detailed diagram of the polyp phase of the model. The focus of this phase is to describe the growth of the polyp to the point where the muscularis mucosa is crossed and the cancer becomes invasive. The assumptions underlying this phase are:

- All cancers arise from an adenomatous polyp;
- The polyp incidence rate applying in the base year also applies throughout the model projection period;
- While polyp incidence varies with age, sex and risk factor groups, the progress of the polyp to invasive cancer is assumed to be independent of these groups;
- Multiple polyps in the same person act independently;
- Polyps which progress to cancer are distributed along the colon and rectum according to the observed cancer distribution but polyps which do not progress to cancer are distributed evenly along the bowel;
- Polyps arise with a diameter less than 10 mm and while they remain at this size they have negligible risk of transition to invasive cancer;
- Polyps remain at their initial size for an average of two years, at which point they either grow or disappear;
- All polyps which grow will ultimately give rise to invasive cancer if not detected first
    - this implies the assumption that only a negligible number of polyps grow without risk of invasive cancer;
- Polyps may only be detected by a screening programme
    - it is possible for polyps to cause symptoms and clinically surface or to be detected during investigation or treatment for other diseases involving the colon, but I will assume the number so detected is negligible;

- Polyps have an upper limit to their growth of a diameter of 75 mm and transition to invasive cancer is assumed to occur at or before this point
    - the size of this upper limit is suggested by the published data on polyp growth.

Invasive cancer

The focus of this phase of the model is the transition of the cancer through the four Dukes' stages and its eventual surfacing, either through screening or clinical detection. This part of the model also assigns a cure status and survival time to each cancer on surfacing. The assumptions underlying this phase are:

- In the absence of detection by screening or clinical surfacing, the disease will arise in stage A and progress through stages B and C to stage D;
- A tumour may clinically surface in stages A to C but will always surface in stage D,
    - in addition, the tumour may be detected by screening in any of these stages;
- There is a single, constant hazard of transition within each stage, leading to a single exponential distribution for the sojourn time;
- At the end of the sojourn time the transition is then governed by the stage specific clinical surfacing or disease transition probabilities and may be either to clinical surfacing or to a subsequent disease stage;
- Screen detection may occur at any time during the sojourn and is governed by the screening model and the stage specific screening sensitivity;
- The sojourn time in and transition from any stage is independent of sojourn times in previous stages and of the polyp phase;
- The cure rate and survival time for those not cured depends only on the stage at diagnosis;
- The progress of the cancer through the four stages is assumed to be independent of age, sex and risk factor status;
- The clinical surfacing probability for a specific stage is independent of the person's age and risk factor status; and
- Multiple cancers in the same person act independently.

3.2.3 Application of the model

The first year for the model projection (the *base year*) is taken to be 1990. The model generates a hypothetical population mimicking the characteristics of the Australian population for that year for the specific age and sex groups under study. This is done using calendar year historical data on the age and sex distribution of the Australian population.

For each calendar year prior to the base year and for each age-sex group, the number of new polyps is calculated using the model's incidence function. The

population units with these polyps are then followed using the model up to the base year. Those assigned to death from colorectal cancer or from a competing cause before the base year are removed. The remainder are taken to be all the members of the population in the base year with an existing polyp or cancer in either the preclinical or clinical state. Finally enough new population units are generated without polyps or cancer to make a total population equal to the actual age-sex group under study in the base year.

This hypothetical population is aged using cause of death probabilities with colorectal cancer deleted to assign a date of death from a cause other than colorectal cancer. Then the model's incidence function is applied to each population unit to generate new polyps for each projected year after the base year. Finally the new and existing polyps and cancer cases are followed to clinical surfacing where applicable and either cure or death from colorectal cancer. Each unit with cancer is taken to have a diagnosed clinical case of cancer if the date of death from another cause is after the date of clinical surfacing. Similarly they are taken to have died from cancer if the date of death from another cause is after the date of death from cancer.

3.3 Previous modelling studies

The first major modelling study of colorectal cancer screening was carried out by Eddy (1980). He used this model to investigate the cost-effectiveness of screening for both an average risk population (Eddy 1984) and a high risk group defined by a family history of colorectal cancer in the absence of a specific genetic syndrome (Eddy et al. 1987, Eddy 1990). Brown also applied Eddy's model to estimate the cost-effectiveness of colorectal cancer screening (Brown 1993) using compliance rates, costs and effects from the randomised trial of faecal occult blood test screening (FOBT) reported by Mandel at al. (1993).

Eddy's modelling strategy was described in chapter two. He uses a time varying Markov framework, but differs from my study in using as his basic assumption that once a disease is detectable by a screening modality then any screen using that modality will always detect the disease. This assumption may apply for screening methods based on visualising the lesion (such as colonoscopy or sigmoidoscopy), where the size of the lesion plays a large part in determining screen sensitivity. However, it is unlikely to apply to screening methods which rely on the detection of blood in the faeces. Bowel cancers and polyps bleed intermittently. The fact that blood may be detected at one point in time does not necessarily imply that bleeding will occur and hence blood be detected at a later point in time.

The effect of breaches of this assumption on the results of Eddy's modelling is unclear. However, Eddy's detailed analysis of the disease natural history does provide a source for some parameter values in my model which cannot otherwise be directly estimated.

Shimbo et al. (1994) studied the cost-effectiveness of colorectal cancer screening in Japan using a similar microsimulation approach to the one used in my study. Their disease model uses parameters largely derived from Eddy's study in a model which

assumes a fixed sojourn time in the polyp phase and each of the cancer stages rather than allowing for variability in these times as my model does. The authors note that the lower mortality from colorectal cancer in Japan, differences in practice patterns and the cost of medical care make their results not directly comparable with studies such as Eddy's for the United States. For similar reasons their results are also unlikely to be directly comparable with my study based on Australian data. However, their detailed analysis of screening characteristics provides a useful input to my discussion of screening in the next chapter.

Wagner et al. (1991) at the US Congress Office of Technology Assessment used a much simpler disease model than ours to investigate the cost-effectiveness of colorectal cancer screening in the elderly. Like Shimbo et al., their disease model assumes a deterministic rather than a stochastic progression. In addition, the assumptions underlying their disease and screening models are intentionally conservative in order to produce cost-effectiveness ratios which represent upper bounds to the true cost-effectiveness. For example, they assume that only 57% of new cancers arise from polyps while noting that there are strong arguments for believing the true value to be higher.

Lieberman (1995) and Salkeld et al. (1996) both developed cost-effectiveness models for colorectal cancer screening without an explicit disease model. Lieberman used the results reported by Mandel et al. (1993) combined with published studies of screening colonoscopy to develop efficacy assumptions for five different screening programmes aimed at average-risk adults using different combinations of FOBT, flexible sigmoidoscopy, colonoscopy and barium enema. Salkeld et al. focused on screening with FOBT for average-risk adults, using results from Mandel et al. to estimate efficacy. From the point of view of my study, Salkeld et al. present the most interesting results because they derive costs for screening and cancer treatment using Australian data. Hence their results are the most directly comparable to ours.

Comprehensive modelling studies of colorectal cancer have been published by Wagner et al. on behalf of the US Congress Office of Technology Assessment (Wagner et al. 1996) and Winawer et al. on behalf of the American Gastroenterological Association (Winawer et al. 1997).

Wagner et al. examined four screening strategies for use among average-risk adults—FOBT, flexible sigmoidoscopy, double contrast barium enema and colonoscopy—individually and in combination. The disease model used is more detailed than that used in their earlier study but still simpler than that used in my study and it still assumes a deterministic rather than a stochastic disease progression. Their model is applied to a synthetic cohort of 100,000 persons and assumes 100% compliance with screening. Thus its results can be seen as the effect of screening on an average person rather than the effect of a population screening programme. So their results are not directly comparable with my study.

Winawer et al. draw on results from Wagner et al. as well as their own simulation study as part of a presentation and justification of a set of clinical guidelines for cancer screening. These guidelines cover both people at average-risk and people in high risk groups. Their simulation study does not have an explicit disease model,

but uses a wide ranging literature search to derive estimates of screening efficacy based on published results. It also does not consider the effect of screening compliance. However, their detailed analysis of screening characteristics provides a useful input to my discussion of screening in the next chapter.

Finally, the MISCAN model described in section 2.6 above has been adapted for colorectal cancer and applied to modelling the cost savings arising from sigmoidoscopy screening at 5 year intervals (Loeve et al. 2000). This modelling work was published after my analyses were completed, so I could not draw on its results. The structure of the model for invasive cancer is similar to mine with the sojourn in each stage based on exponential distributions with similar mean times to those I used. However, the adenoma phase is modelled using slightly different assumptions to my model.

3.4 Data sources for model parameters

3.4.1 National Cancer Statistics Clearing House

The National Cancer Statistics Clearing House (NCSCH) is operated by the Australian Institute of Health and Welfare under the supervision of the Australasian Association of Cancer Registries. It provides a facility for compiling data produced by individual State and Territory cancer registries on a continuing basis, identifying multiple registrations and producing accurate national incidence statistics.

The NCSCH receives data from individual State and Territory cancer registries on cancers diagnosed in usual residents of Australia. The collection commenced with cases first diagnosed in 1982. The most recent year for which data had been collected from all States and Territories at the time this model was developed was 1990, though the collection has since been extended to 1996. The data cover details of both the cancer and the person with the cancer. However, stage at diagnosis is not one of the variables collected.

The data set extracted from the NCSCH and used in this study covers all usual residents of Australia who were diagnosed with colorectal cancer in the period 1982 to 1990. The data includes the age at diagnosis, sex and state of usual residence of the person and the site of their first primary cancer. These data were used as the basis for estimating the model's incidence function.

3.4.2 South Australian Hospital registry data

The estimates of stage specific clinical surfacing probabilities and stage specific survival rates are based on data collected in South Australia by the hospital based cancer registries. These data cover patients treated in three Adelaide teaching hospitals—the Flinders Medical Centre, the Queen Elizabeth Hospital and the Royal Adelaide Hospital.

Data collection for the hospital based registries commenced in 1987 but included records for all patients treated in these hospitals from 1980. The data set includes data up to the first part of 1996. Each record covers one cancer and includes the cancer's date of diagnosis, site of primary cancer (coded according to the ICD-9 classification) and stage at diagnosis and the patient's sex, date of birth and, if applicable, date and cause of death. The cause of death is coded according to the ICD-9 classification if it is a cancer but given the dummy coding 9999 for any other cause.

Death has been determined by matching the patient records against both the state register of deaths and the national death index. Since registration of deaths in Australia is effectively complete and the matching process highly effective, this means that most patients who died in Australia during the collection period have had their death recorded in the register. However, some will have been missed in the matching process and some will have left Australia before death and so not be on the register. The analysis described below takes account of this loss to follow-up.

A small number of these records cover multiple cancers in the one patient. The records do not identify the patients, but probable multiple records for a single patient can be identified by matching the records on sex, date of birth and, where applicable, date and cause of death. Among the 1348 records for people who were still alive, 55 records were identified where two or more records shared the same sex and date of birth. Among the 1943 records for people who had died, 10 records were identified where two or more records shared the same date of birth, sex, date of death and cause of death.

For each of the probable matches, the record with the earliest date of diagnosis was taken as the first clinical surfacing of the cancer. These were combined with the records for people with a single cancer and the resulting data set used to estimate the probability of clinical surfacing. Thirteen of these probable matches had two cancers with the same date of diagnosis but different stages. These were excluded from this analysis because it was impossible to identify which cancer was responsible for the clinical surfacing.

Ninety nine records, including two of the probable matches, had dates of diagnosis before 1980. These represent patients diagnosed before the period covered by the registries who were still receiving treatment during the registry period. The fact that they survived at least till the beginning of the registry period suggests a less serious form of the disease. Hence they have been excluded from the analysis as representing a possible bias in both the estimates of the surfacing probabilities and survival times.

A further 225 cancers had no recorded stage at diagnosis and so were also excluded from the analysis.

Unlike the National Cancer Statistics Clearing House data, which cover all cases of cancer diagnosed in Australia, the hospital based registry data are not collected to be representative of cancer in the general population. Cases attending the South Australian teaching hospitals tend to come from the less affluent and older sectors

of the community and reflect a variety of referral patterns (South Australian Health Commission & South Australian Cancer Registry 1994).

Figures 3.7 and 3.8 show the differences in age and sex composition between the hospital based registry cancers and those not on the hospital registry but included on the South Australian state cancer registry. These show the hospital register as being generally older and comprising a higher proportion of males and a lower proportion of females. Stage is not recorded on the state register so a direct comparison by stage is impossible. However, it is likely that patients on the hospital register present with a more advanced form of the disease, so results based on this register are likely to be biased. Unfortunately there are no population based cancer registers which record stage at diagnosis, so the hospital based register remains the best available data source for estimates of stage distribution at clinical surfacing and stage specific survival rates.

3.4.3 General demographic data

National mortality data

The model requires an estimate of the probability of death in the absence of colorectal cancer. This is calculated using the technique described by Chiang (Chiang 1968). Let $\hat{p}$ be an estimate of the probability that an individual of a particular age and sex who is alive at the beginning of the year survives to the end of the year (Omitting subscripts for age and sex for clarity). Further, let D and D_c be the total number of deaths in the base year from all causes and from colorectal cancer respectively. Then the probability that an individual who is alive at the beginning of the year will die during the year if colorectal cancer is eliminated as a cause of death is estimated by

$$\hat{q}_{\bullet c} = 1 - \hat{p}^{D_c/D}$$

National mortality data are collected by the Registrars of Births, Deaths and Marriages in each State and Territory, compiled into national data by the Australian Bureau of Statistics and then provided to the Australian Institute of Health and Welfare. Counts for D and D_c were extracted from the Institute's mortality database for 1990 (the model's base year). The estimate $\hat{p}$ was taken from the Australian life tables for the period 1990 to 1992 as published by the Office of the Australian Government Actuary (1995).

National population data

The model also requires population data for the base year and historical data for preceding calendar years. These were taken from the Australian Bureau of Statistics estimates of the total mid-year Australian population.

3.4.4 Published data on polyp growth

Published studies are the main source of data on polyp size, growth and transition to invasive cancer. There are a number of these studies, but not all of them present data in a form which is useful for my model.

Hoff et al. present a study of changes over a two year period in people with small polyps (Hoff et al. 1986). Their study was part of an endoscopic population screening study for colorectal polyps among 200 men and 200 women aged 50 to 59 years. Polyps 5 mm or larger in diameter found on initial colonoscopic examination were removed but smaller polyps had their location and size recorded and were left unresected. Subjects with these smaller polyps were offered a follow up colonoscopic examination after an interval of two years.

The results of this study support the proposition that polyps may either grow, remain stable in size or shrink and disappear. However, the data are not presented in such a way as to allow any estimates of initial size distribution or growth rate.

Stryker et al. present a retrospective study derived from records of the Mayo Clinic from a six year period just before the advent of colonoscopy. This identified 226 patients with colonic polyps greater than 10 mm in diameter in whom periodic radiographic examination was elected over excisional therapy (Stryker et al. 1987).

This study demonstrates growth in polyps of this size and their potential for development of invasive cancer. However, it also does not present data in such a way as to allow any estimates of growth rate.

Knoernschild studied 257 patients with asymptomatic benign polyps detected through routine sigmoidoscopy who agreed to allow the lesion to be re-examined at six to 12 month intervals. The polyps ranged in size from two to 15 mm in diameter, and those that grew were removed when they attained a 15 mm diameter (Knoernschild 1963).

This study supports the proposition that small polyps may grow into larger polyps with a risk of malignant transformation, but does not present any data which would allow model parameter estimation.

The two most useful studies from the point of view of determining parameters for my model were conducted by Figiel et al. (1965) and Welin et al. (1963).

Figiel et al. present a retrospective study of 18 patients with colorectal cancer who had previous radiographic studies of non-invasive polyps. This allowed the observation of the growth of the polyps prior to malignant transformation. The authors calculated the growth rates of this series of polyps. In addition they present a summary review of a further 300 cases in which polyps were observed periodically.

Welin et al. followed the growth of 259 polyps by repeated roentgenographic examinations. Twenty of these ultimately developed invasive cancer, visible as marginal infiltration, and were removed. They present data on the initial size of polyps, their growth rates and their size at malignant transformation.

3.5 Parameters for the invasive cancer model

3.5.1 Sojourn times and transition probabilities for each cancer stage

Once the polyp has made the transition to invasive cancer, the cancer passes through some or all of the four stages. The information required for this part of the model is:

- the probability distribution governing the sojourn time in each stage; and
- the probability which governs the transition out of that stage to either clinical surfacing or the next stage.

Sojourn times

In principle each preclinical cancer stage has two sojourn time distributions—the sojourn time before transition to the next cancer stage and the sojourn time before clinical surfacing. There are insufficient data to model these sojourn times separately. Instead this model will follow the approach used by van Oortmarssen et al. in their breast cancer model (van Oortmarssen et al. 1990) where each stage has a single constant hazard of transition, leading to a single exponential distribution for the sojourn time. The transition is then governed by the stage specific clinical surfacing or disease transition probabilities and may be either to clinical surfacing or to a subsequent disease stage.

Screen detection may occur at any time during the sojourn and is governed by the screening model and the stage specific screening sensitivity.

The exponential distribution is fully specified by its mean. In principle this mean may be modelled from data derived from randomised controlled trials of screening. Unfortunately the published data from the colorectal cancer screening RCT's are not sufficiently detailed for this purpose.

Eddy derived estimates of the mean sojourn time in stages A, B and C through consultation with a panel of experts with wide experience with colorectal cancer (Eddy et al. 1987). Although these estimates have no empirical basis, they have been widely used in modelling studies of colorectal cancer by Eddy (Eddy 1984, Eddy 1990, Eddy et al. 1987) and others (Shimbo et al. 1994, Brown 1993). In the absence of better estimates, I will also use them. Further, I will assume that they apply regardless of the cancer site and the patient's age and sex.

Eddy did not consider stage D cancer in his model and so provided no estimate for its sojourn time. Stage D is regarded as incurable, so screen detection of cancers in this stage confers no benefit. It does, however, add to the cost of a screening programme and so should be included in the model.

Stage D cancer is associated with a short survival time. Further, given that stage D cancer involves distant metastases, one may assume a rapid development of symptoms. Hence the average sojourn time is likely to be short compared with that for the other stages. I have arbitrarily taken the mean sojourn time as 0.75 years.

The true sojourn time in stage D is unlikely to be longer than that for stage C, so this value is unlikely to be a gross under estimate. It may over estimate the true sojourn time but if it does so the main effect will be to inflate the cost of screening relative to its benefits. The size of any such effect will be investigated by subjecting the mean sojourn times to sensitivity testing.

Table 3.5 presents the mean sojourn times used in my model.

Probability of clinical surfacing

Age adjustment

Clinical surfacing will be influenced by both the appearance of symptoms and willingness to seek treatment for those symptoms. This model will follow other colorectal cancer models in assuming that the appearance of symptoms is independent of age and sex—that is the disease in a specific stage at a specific site is equally likely to present symptoms at any age in both sexes. Women in Australia are statistically more likely to use medical services than men, so it would be unsafe to assume that men and women are equally likely to seek treatment for symptoms. Thus the clinical surfacing distribution may differ between sexes. However I will assume that in the age ranges of interest to screening there is no difference in willingness to seek treatment by age.

The frequency distribution of cancer stages at diagnosis at any specific age will depend on both the age-specific prevalence of cancers in each stage and the probability of clinical surfacing in each stage. Hence, although I am assuming no difference in stage specific clinical surfacing probability by age, the observed age specific rates will vary. The variables sex, age and site of cancer are common to both the hospital based registry and the national clearing house data set. Hence an 'age adjusted' estimate of the probability of surfacing at each stage can be calculated by applying standard demographic age adjustment techniques. This involves calculating the surfacing probability for each stage in each age, sex and site group for the hospital register, assuming that this holds for the equivalent group for the national data set and then calculating an overall probability for each stage as the weighted sum of probabilities across all ages. If the observed rate of surfacing in stage j at age a for sex s and site i is p_{jasi} calculated from the hospital registry data then

$$p_{jsi} = \sum_a \left(p_{jasi} * n_{asi} \right) \Big/ \sum_a n_{asi}$$

where

n_{asi} = the number of people on the national clearing house data set with age a, sex s and site i.

is the estimate of the overall probability of surfacing at stage j for sex s and site i. This adjustment allows for the different age distribution between the two data sets but does not allow for differences in the disease severity within each age group. It also does not allow for other possible differences such as differences by place of residence which may be due to accessibility of treatment services.

Stability over time

It would be useful to be able to pool all the hospital registry data to calculate age, sex, site and stage specific surfacing rates, but I can only do this if the observed rates are stable over time. This was investigated using a log linear modelling approach with the number of cancers as dependent variable and age, sex, year of diagnosis and stage at diagnosis as covariates. In this model the main effects for all four covariates and all the interaction terms for age, sex and year of diagnosis are treated as 'nuisance' parameters—they are of no interest in themselves but must be included in the model. The effect of each of age, sex and year on stage is then measured by the relevant interaction terms with the stage effect.

Table 3.6 lists the number of people with cancer recorded on the hospital based registry classified by year of diagnosis, primary site, sex, age at diagnosis and stage at diagnosis. Year of diagnosis has been classified into two groups—1980–89 and 1990–96. Age at diagnosis has been classified into five groups—Under 50 years, 50–59, 60–69, 70–79 and 80 and over.

Standard log linear modelling makes no assumptions about the ordering of categories within any of the variables. This is obviously not an issue for the variables with only two categories—year and sex. However, modelling should take account of possible ordering in the categories in the other variables. This was done using row models, column models and uniform association models as described by Agresti (1990). These may be regarded as forms of log linear model and hence may be incorporated into a log linear analysis.

Table 3.7 lists the results of fitting each of stage at diagnosis, primary cancer site and age at diagnosis as ordinal variables. The deviance statistic and associated degrees of freedom may used as a test of the goodness of fit of each of these variables as compared with the fully saturated model which makes no assumption about ordinality. These indicate that while site and age may be regarded as ordinal, treating stage as ordinal results in a significantly poorer fit.

Table 3.8 is an analysis of deviance table presenting the results of modelling stage as a function of age, sex, site and year. These demonstrate that, while age, sex and site have significant effects on stage, the effect of year on stage may be removed from the model without significant loss of fit. Hence I can assume stability over time in estimating clinical surfacing probabilities.

Table 3.9 presents the age adjusted estimates of the rates of clinical surfacing by sex and primary site. These are used in the model as the probability that a specific cancer will surface at the end of its sojourn in each stage.

3.5.2 Survival times and cure rates

The model assumes that the prognosis of a person with a diagnosed colorectal cancer depends entirely on the stage at diagnosis irrespective of whether the cancer was detected by screening or surfaced clinically. A proportion of people in stages A, B and C will be completely cured and the rest will die from the disease unless they die from another cause first. Stage D colorectal cancer is assumed to be incurable.

Death from another cause is assumed to operate independently of whether or not the person has cancer. The proportion cured and the time to death from colorectal cancer for those not cured is assumed to be independent of age and sex.

The hospital based registry data were used for estimating the proportion cured and identifying a probability distribution for the time to death from colorectal cancer. This introduces a further complicating factor. The occurrence of death and its date and cause are determined by matching the registry against the State and national deaths registers. No such matching process is perfect, so there will be a proportion of registrants who have died but who were not identified in the matching and whose deaths will not be reported. There will also be a proportion of registrants who have left the country after diagnosis and whose deaths, if they occur, will not be recorded on the deaths register. Both of these groups will appear to have an infinite survival time. Unless the analysis of survival times is adjusted for this loss to follow-up, the results may be severely biased.

Tallis et al. (1988) have discussed the problem of analysis of survival data from a registry with passive follow-up and the analysis presented here has been adapted from their work. Their analysis focused on death from any cause. The analysis presented here differentiates between death from cancer and death from another cause and includes an estimate of the proportion cured of cancer.

Definitions, assumptions and estimating equations

I will start with the assumption that the proportion of registrants lost to follow-up is constant and not related to age, sex, disease stage, time of entry onto the register or length of time on the register. This is likely to be true for those missed in the matching process but may not be true in general. However, Tallis et al. suggest that for losses of up to 10% to 15% this simple model provides a good approximation to more complex set-ups.

Let

r	=	the age of the registry (in years);
$p_1, p_2, \ldots, p_r$	=	the yearly proportion of the total registrants;
θ	=	the proportion of registrants lost to follow-up;
c_A, c_B, c_C, c_D	=	the proportion cured in stages A, B, C and D (which may be 0);
Y	=	the survival time till death from colorectal cancer in the absence of other causes of death for those people not cured.

I will first consider only deaths from colorectal cancer. Consider individuals entering the register in the j^{th} year. They will either die in their first, second,… or $(r+1-j)^{th}$ year on the register with probability $P_1, P_2, \ldots, P_{r+1-j}$ or they will survive to the end of the registry period with probability P_j^*.

Next I will consider deaths from other causes. In principle I could use this information to model competing causes of death directly. Unfortunately there are not enough observations to do this. Instead I will assume independence between

death from colorectal cancer and death from other causes. The fact that an individual died from another cause in their ith year on the register means that they did not die from colorectal cancer in their first i-1 years. For those who would have otherwise died from colorectal cancer at a later time the probability of this is denoted P_i^{**}.

Now I will define the following events:

E_i is the event of an observed death from colorectal cancer in an individual's ith year on the register;

E_i^{**} is the event of an observed death from another cause in an individual's ith year on the register;

E_j^{*} is the event of an individual who entered the register in its jth year and who was not observed to die from any cause before the end of the registry period (ie an observed survivor).

For an individual diagnosed with colorectal cancer in stage X (where X = A, B, C or D) in the registry's jth year, the events $E_1, E_2, \ldots, E_{r+1-j}$ have the probabilities
$(1-c_X)(1-\theta)P_1, (1-c_X)(1-\theta)P_2, \ldots, (1-c_X)(1-\theta)P_{r+1-j}$;
the events $E_1^{**}, E_2^{**}, \ldots, E_{r+1-j}^{**}$ have the probabilities
$(1-\theta)\left(c_X + (1-c_X)P_1^{**}\right), (1-\theta)\left(c_X + (1-c_X)P_2^{**}\right), \ldots, (1-\theta)\left(c_X + (1-c_X)P_{r+1-j}^{**}\right)$;
and the events $E_1^{*}, E_2^{*}, \ldots, E_{r+1-j}^{*}$ have the probabilities
$\theta + c_X(1-\theta) + (1-c_X)(1-\theta)P_1^{*}, \theta + c_X(1-\theta) + (1-c_X)(1-\theta)P_2^{*}, \ldots,$
$\theta + c_X(1-\theta) + (1-c_X)(1-\theta)P_{r+1-j}^{*}$.

I can express the probabilities P_i, P_i^{**} and P_j^{*} in terms of the underlying distribution function governing time to death from colorectal cancer $F(y)$. Assuming that reporting is uniform within each reporting year, then

$P_i = \int_0^1 \left[F(i-l) - F(i-1-l)\right]dl$;

$P_j^{*} = 1 - \int_0^1 F\left[r - (j+1) - l\right]dl$; and

$P_i^{**} = 1 - \int_0^1 F\left[(i-1) - l\right]dl$

The probability of event E_i over all reporting times is

$$Q_i = (1-c_X)(1-\theta)P_i \sum_{j=1}^{r+1-i} p_j = (1-c_X)(1-\theta)P_i\Phi_i .$$

Similarly the probability of event E_i^{**} over all reporting times is

$$Q_i^{**} = \left(c_X + (1-c_X)(1-\theta)P_i^{**}\right)\Phi .$$

Finally the probability of event E_j^* for a given stage X is

$$Q_j^* = p_j\left(\theta + c_X(1-\theta) + (1-c_X)(1-\theta)P_j^*\right).$$

Now considering the observed deaths and observed survivors, events $E_1, E_2, \ldots, E_r$, $E_1^*, E_2^*, \ldots, E_r^*$, $E_1^{**}, E_2^{**}, \ldots, E_r^{**}$ we obtain the log likelihood

$$\log L = \sum_{i=1}^{r} N_i \log Q_i + \sum_{i=1}^{r} N_i^* \log Q_i^* + \sum_{i=1}^{r} N_i^{**} \log Q_i^{**}. \quad (1)$$

Where N_i, N_i^*, N_i^{**} are the observed frequencies of E_i, E_i^*, E_i^{**} respectively.

If I parameterise *F* as $F(y;\gamma)$, then in principle I can maximise this function for the unknown parameters θ, c_X and γ. The two difficulties with this in practice are (1) distinguishing between a registrant lost to follow-up and a person cured of cancer (since both mainly appear as censored observations at the end of the registry period); and (2) distinguishing between people with a very long survival time (such as is found in stages A and B) and those who are either lost to follow-up or cured. The way I will approach this is to use a two stage estimation—I will first estimate γ and then maximise the log likelihood in equation 1 over θ and c_X while holding γ constant at its estimated value. If we consider only events $E_1, E_2, \ldots, E_r$ (the observed deaths from colorectal cancer), then the appropriate log-likelihood, which is a function of γ only, is

$$\log L = \sum_{i=1}^{r} N_i \log\left(Q_i / Q_\bullet\right), \text{ where } Q_\bullet = \sum_{i=1}^{r} Q_i. \quad (2)$$

Maximising this gives me an estimate of γ.

Since stage D colorectal cancer is regarded as incurable, I can set $c_D = 0$ allowing me to maximise the log likelihood in equation 1 for stage D, holding γ constant at its estimated value, to give an estimate of θ. This estimate is then used in the log-likelihood equations for stages A, B and C to give estimates of c_A, c_B and c_C respectively.

Parameter estimates

The data used for estimating the survival time distributions and the cure rates are presented in tables 3.10 and 3.11. The analysis was done using the Splus statistical analysis software (Mathsoft, 1995). The log-likelihood functions were maximised using a combination of searching over a grid of parameter values and the Splus function *minfun* . The integrals within the log-likelihood functions were evaluated using the Splus function *gkint*.

Three forms were tried for the underlying survival distribution *F*—exponential, Pareto and lognormal. The exponential distribution is commonly used to model the time to an event. It is specified by the single parameter λ.

The Pareto distribution may be regarded as a mixture of exponential distributions and hence would be appropriate for a group which is heterogeneous with respect to

factors possibly influencing survival such as age at reporting and treatment undertaken. The distribution function may be written

$$F(y;\delta,k) = 1 - \left(\frac{\delta}{y+\delta}\right)^k$$

This distribution is over dispersed relative to the exponential distribution and for small k has a long tail so it can accommodate long survivors (Tallis et al. 1988).

The lognormal distribution has been successfully used for modelling survival times among breast cancer patients (van Oortmarssen et al. 1990). In this case the survival time is modelled as the log of a normal distribution with mean μ and standard deviation σ.

The estimates of the distribution parameters are presented in table 3.12. Table 3.13 presents a summary of chi-square goodness of fit tests comparing fitted values from each of the hypothesised distributions with the observed deaths data. The fit for each of the distributions is fairly good for short survival times. The differences between them arise from how well they fit the longer survival times. All three distributions provided a good fit for stages A and B. The exponential was a poor fit to stages C and D and the Pareto was a poor fit to stage C. The lognormal distribution was a good fit for stage D and a moderate fit for stage C.

The lognormal distribution provides the best fit across all the disease stages, so it was chosen as the basis for estimating the proportion lost to follow-up and the cure rates for each stage. These estimates are presented in table 3.14.

The parameter estimates for cure rates given in table 3.14 appear lower than expected, particularly for stages A and B. For example, a South Australian analysis of five year survival by stage gave an estimate of close to 90% for stage A (South Australian Health Commission & South Australian Cancer Registry 1994). The small number of deaths from stage A cancer after five years suggests that an ultimate survival rate of 64% is unlikely. The reason for the underestimation is the fact that my estimation procedure did not differentiate well between people cured of cancer and those who ultimately die but have a long survival time. The estimates of the log normal parameters μ and σ were based on the observed deaths and the estimate of the loss to follow-up rate θ was based on the stage D cancers which all have a relatively short survival time. Hence these estimates were not affected by the confounding between long survival and cure.

Another analysis of cure rates was done using the SAS PROC LIFETEST procedure which estimates survival rates using a non-parametric method. These estimates will not take account of the loss to follow-up, but I can adjust them using the estimated rate of loss to follow-up θ. The results are given in table 3.15. These higher estimated cure rates appear more plausible, particularly for stages A and B. The LIFETEST procedure also gave an estimate of five year survival rate for stage A cancers of 88%, which compares well with the South Australian analysis.

I will use the LIFETEST procedure cure rate estimates in my model but investigate the effect of using the lower cure rate estimates in the sensitivity testing.

3.6 Parameters for the polyp-cancer sequence

3.6.1 Proportion of cancer originating as polyps

As noted above, the adenoma-cancer sequence is almost universally accepted for the majority of colorectal cancers. However, the proportion of cancers originating in this way remains a subject of some debate.

Eddy's first modelling study used arbitrary estimates of 75% and 25% as trial values for this proportion (Eddy 1980). However, his later models use the much higher figure of 93% (Eddy et al. 1987, Eddy 1990). Wagner et al. use the (admittedly conservative) figure of 57% in their first model (Wagner et al. 1991) and 70% in their later model (Wagner et al. 1996), though with a sensitivity analysis using 56% and 90% as upper and lower limits. Shimbo et al. quote Wagner et al.'s figure of 57% and a figure of 80% which they mistakenly attribute to Eddy's 1990 study as partial justification of their own even more conservative assumption of 50% (Shimbo et al. 1994).

Morson (1974), Muto et al (1975) and Fenoglio and Lane (1974) present strong arguments that, apart from the case of predisposing causes such as ulcerative colitis, cancers do not arise 'de novo' without a prior polyp phase. This implies that in general such de novo cancers make up less than 1% of all colorectal cancer. Further, since cancer associated with inflammatory bowel disease is beyond the scope of this study, such de novo cancers would play no part in my modelling.

I will follow this argument by assuming that all colorectal cancers arise from adenomas.

3.6.2 Initial size distribution

Welin et al (1963) present a plot of the diameters of 239 polyps with no evidence of invasive cancer, plotted on a log scale. These show a roughly bell-shaped curve centred at five mm and ranging between two mm and slightly greater than 10 mm. Hence a lognormal distribution was assumed for these initial sizes, with a mean at five mm and with 99% of tumours falling in the range 2.5 mm and 10 mm. This log normal distribution will be denoted $l(i)$, where i is the initial polyp diameter.

3.6.3 Size distribution at transition to invasive cancer

Most authorities agree that invasive cancer is very rare in polyps under 10 mm in diameter (Muto et al. 1975)—hence although such cases exist (Welin et al. 1963) I will assume that their prevalence is negligible for the purposes of this model. Polyp size also has an upper limit determined by the polyp reaching the limits of nutrition (Welin et al. 1963).

Welin et al. followed the growth of 259 polyps by repeated roentgenographic examinations. Twenty of these ultimately developed invasive cancer, visible as marginal infiltration, and were removed. The actual transition to cancer would have occurred in the approximately 12 months between examinations, so there may have been further polyp growth after the transition. However, the very slow rates of growth ensure that the measurements of polyp size at the final examination were close enough to the size at transition for the purposes of this model. Figure 14 of Welin et al. gives a distribution of these polyp diameters at the final examination, reproduced in Table 3.16 below. These data suggest an upper limit in polyp diameter of around 75 mm.

The risk of invasive cancer increases with the size of the polyp (Muto et al. 1975). Hence the model of the size distribution of polyps at transition to invasive cancer was based on the assumption of the hazard of transition being proportional to the size of the polyp. This suggests the use of a Weibull distribution. The hazard function can be written as

$$w(d) = \lambda\gamma d^{\gamma-1} \tag{3}$$

where d is the polyp diameter. If I take the polyp diameter as the measure of size, then $\gamma = 2$ in equation 3. If I take the polyp volume as the measure of size, then $\gamma = 3$. This arises from the fact that colonic adenomas tend to grow as spreading plaques with a roughly constant thickness approaching the shape of cylinders (Welinet al. 1963). Hence the volume is proportional to the square of the diameter.

The parameters for the fitted Weibull distribution, conditional on $10 \le d \le 75$, were found by holding the γ value at 2 and 3 and varying the λ parameter to give the closest fit to the data in Table 3.16 (using the χ^2 value as the measure of goodness of fit). Finally both λ and γ were varied to achieve the closest fit to the data. The results are given in Table 3.17.

Although each of these is an adequate fit, the fit for $\gamma = 3$ is a marked improvement over $\gamma = 2$. The estimated value of γ is close to 3 and using it does not improve the fit. So the Weibull distribution with parameters $\gamma = 3$ and $\lambda = 0.000015$ was used to model the polyp size distribution at transition to invasive cancer.

3.6.4 Polyp initiation and growth rates

Non cancerous polyps

A large proportion of polyps never grow and have no clinical significance. The detection of these by screening incurs a cost but confers no benefit. Their occurrence is assumed to be uniformly distributed throughout the colon and they are assumed to be under 10 mm in diameter and stable in size (Muto et al. 1975, Welin et al. 1963). There are no direct observations of their duration, but Eddy argues that they must have an average duration of about two years (Eddy 1980).

I will use the figures reported in Winawer et al. (1997) and assume that 60% of men and 40% of women will have an adenomatous polyp at some time in their life. This does not account for other types of polyps, but I will assume that their impact on

screening is negligible. Further, I will assume that each person has at most one of these non cancerous polyps and that polyp initiation occurs after the age of 40 with a constant risk throughout the rest of a person's life. Most people with an adenomatous polyp (either cancerous or non cancerous) have only one such polyp (Nicholson et. al. 2000). Further, the size of a polyp is directly related both to the probability of high-grade dysplasia in that polyp and to the probability of other adenomatous polyps elsewhere in the colon and rectum (Winawer et. al. 1997). Hence although there are no data to directly support this assumption, it is consistent with what is known about polyp distribution. When a polyp arises its duration is exponentially distributed with a mean of two years.

Cancerous polyp initiation

A plot of the incidence of colorectal cancer against age shows the sigmoid shaped curve which is characteristic of most cancers in Australia, being small for young people, increasing fairly sharply after the age of 50 and then levelling off (figure 3.1). Since the vast majority of these cancers arise from polyps, it seems reasonable to assume that the rate of polyp initiation will follow a similar curve. Hence the initiation rate *h(a)* will be modelled as a function of age *a* thus:

$$h(a) = \frac{c_0 e^{c_1 + c_2 a}}{\left(1 + e^{c_1 + c_2 a}\right)} \tag{4}$$

Here the parameters c_1 and c_2 model the shape of the curve while c_0 allows the curve to approach a maximum value less than 1.

Cancerous polyp growth

Both Welin et al. (1963) and Figiel et al. (1965) present measurements of growth rates for colonic polyps estimated from serial roentgenologic observations. These are not presented in a form which can be used for directly estimating the distribution of growth rates, but they do provide some useful insights into polyp growth.

Welin et al. observe that the mechanics of neoplastic growth imply that unrestrained cellular proliferation leads to an exponential growth in tumour volume. However, it is possible that other variables could alter the exponential growth and produce a variety of growth patterns. The data they present are consistent with steady growth but could not distinguish between a linear or an exponential growth pattern.

Figiel et al measured growth rates for 18 polyps which ultimately developed invasive cancer. Fifteen of these were based on 2 size measurements and only 3 were based on 3 size measurements. However they note that these 3 were in close agreement with the expected sizes if the growth was exponential.

Welin et al observe that, assuming exponential growth, the diametric growth for cancers is in the range 0.0003 to 0.0025 mm per day. The data presented by Figiel et al are in the range 0.0007 to 0.0030 mm per day.

The model developed here assumes steady exponential growth with diametric growth rates in the range 0.0003 to 0.0030 mm per day. This can be written as

$$d = d_0 e^{\beta t} \tag{5}$$

where d_0 is the diameter obtained from the lognormal distribution $l(\bullet)$ described above.

The distribution of growth rates are assumed to follow a modified Beta distribution with two parameters and a range adjusted to conform with the observed range. Specifically a growth rate β is modelled as:

$$\beta = 0.0003 + (0.003 - 0.0003)W \tag{6}$$

where W has the density function

$$f_W(w) = \frac{w^{p-1}(1-w)^{q-1}}{B(p,q)} \tag{7}$$

and where $B(p,q)$ is the complete Beta function with parameters p and q. This is similar to the approach used by O'Neill et al. (1995) in modelling the growth of breast tumours.

Estimating parameters for polyp initiation and growth

The model for polyp initiation and growth requires estimates of the parameters in equations 4 and 7. These estimates were derived following the procedure described by O'Neill et al. (1995). The expected frequency of new cases of cancer was calculated using initial guesses for these parameters and the invasive cancer model described above. This expected frequency was then compared to the actual numbers of new cancers from the national cancer registry data. This process was iterated for different values of the parameters until a satisfactory fit was found using minimum χ^2 criteria.

The expected frequency was calculated from calendar year historical data on the age distribution of the Australian population. For each age by calendar year group, the number of new polyps which would eventually lead to a cancer was calculated using the initial guesses of the parameter values of $h(y)$. These polyps were assigned an initial size at random from the lognormal distribution $l(i)$ and a growth rate at random from the distribution $f(\beta)$ using initial guesses of its parameters. The polyps made the transition to invasive cancer according to the Weibull distribution $w(d)$ and surfaced after a time governed by the sojourn time and clinical surfacing distributions described above.

The population members with the polyps were also subject to mortality from other causes. They were each followed until their cancer surfaced or they died from another cause, which ever happened first. The cancers surfacing in the estimation period were counted and compared with the actual numbers of new cancers.

Colorectal cancer incidence has fluctuated somewhat over the period covered by the national cancer registry, but the plot of the data in figure 3.2 shows that it has remained relatively stable over the four years 1987 to 1990. The initiation and

growth parameters were estimated using data for the period 1987 to 1989 and tested by comparing the model's projected values for 1990 with the observed values.

This procedure was carried out separately for men and women and for each primary site. The results of the model fitting are presented in table 3.18 and the parameter estimates are presented in table 3.19. These show a reasonably good fit between the model and the observed data.

The expected frequencies and associated χ^2 values depend on the particular random allocation of growth rates, sojourn times and surfacing distributions from the model. Another run of the procedure would give rise to a different random allocation leading to different frequencies and a different χ^2 value. Since the fitting procedure is designed to find the smallest χ^2 value, other χ^2 values for the same set of parameter values are likely to be higher than the ones presented in the table. A further check on the model which takes account of this variation is to simulate a confidence interval for the model projections and compare this with the observed values.

The model was used with the estimated parameter values to project 199 separate frequency distributions for new cases of cancer for the period 1987 to 1989. If the model is consistent with the observed data, then the projected and observed frequencies together should constitute 200 observations from the same probability distribution. Hence when the counts for each age group are ordered from smallest to largest, the second smallest and the second largest constitute approximate 99% confidence bounds and the sixth smallest and sixth largest constitute approximate 95% confidence bounds for that age group.

Figures 3.9 and 3.10 are a plot of these confidence bounds along with the observed data for all new cancer cases in the period 1987 to 1989 for men and women. The observed counts for women are within the 95% confidence bounds for all age groups. The observed values for men lie within the 95% confidence bounds for all ages except 65–69, where it is just below the lower 95% confidence bound.

Figures 3.11 and 3.12 are a plot of the simulated confidence bounds along with the observed data for all new colorectal cancer cases in 1990 for men and women. The observed counts for both men and women are within the 95% confidence bounds for all age groups except for ages 50–54, where they are within the 99% confidence bounds. This suggests that the model projections are consistent with the observed data.

Synchronous and metachronous cancers

People with one adenomatous polyp are at an increased risk of developing another polyp, either at the same time as the first polyp (synchronous) or at a later time (metachronous). For the purposes of this study, we will define a synchronous polyp as one which arises in the same year as the first polyp and a metachronous polyp as one which arises in a subsequent year.

Cali et al investigated the risk of metachronous cancer in a retrospective analysis of 5476 average-risk individuals entered in a tumour registry (Cali et al. 1993). They found that this risk was constant over 20 years from the first cancer. However, their

published risk estimate only covers invasive cancers and so is an under estimate for the risk of metachronous polyps.

Ahlquist et al conducted a prospective study of 1,217 patients who had undergone resection of colorectal cancer (Ahlquist et al. 1993). The main aim of the study was to assess the accuracy of FOBT, but as part of the study these patients underwent annual endoscopic and/or radiographic evaluations for up to the first three postoperative years. Patients found to have colorectal cancer did not continue in the study.

Ahlquist et al report that of a total of 2,293 annual examinations, 386 found a single polyp, 16 found a polyp and a new cancer, one found a new cancer only, one found a new cancer and a recurrence of the existing cancer and 28 found recurrence of the existing cancer only.

Given the slow progression of the polyp cancer sequence, it seems unlikely that the new cancers arose from polyps which also arose in the examination year. The most likely explanation is that these lesions were present before the examinations began. Hence, if I ignore recurrence, then I can estimate the probability of at least one polyp developing in any year as

$$\frac{386+16}{2293} = 0.17$$

None of the examinations found two or more polyps, so I will assume that at most one polyp develops in any year.

Further, Ahlquist et al report that 14% of the new polyps were greater than 10 mm in diameter. If I assume that these are the ones that will progress to cancer, then I can use this as my estimate of the proportion of metachronous polyps which will progress to cancer. One difficulty with these estimates is the assumption that none of the polyps under 10 mm would ever grow and progress to cancer. Violation of this assumption would lead to my estimate of the malignancy potential of metachronous polyps being an under estimate. This will be subject to sensitivity testing in the final model.

3.7 Tables

Table 3.1: Most frequently occurring cancers in Australia, 1996

Cancers	New cases			Deaths		
	Number	Per cent of all new cancer cases	Lifetime risk	Number	Per cent of all cancer deaths	Potential years of life lost
Colorectal	10,998	14.2	1 in 21	4,606	13.5	30,903
Prostate	10,055	12.9	1 in 21	2,644	7.8	6,228
Breast	9,706	12.5	1 in 23	2,640	7.7	31,143
Skin-melanoma	7,761	10.0	1 in 29	903	2.6	10,775
Lung	7,621	9.8	1 in 28	6,764	19.8	46,020
Non-Hodgkin's lymphoma	3,105	4.0	1 in 73	1,388	4.1	11,885
Unknown primary site	3.031	3.9	1 in 87	2,331	6.8	13,678
Bladder	2,544	3.3	1 in 91	778	2.3	2,745
Kidney	2,000	2.6	1 in 109	793	2.3	5,788
Stomach	1,857	2.4	1 in 129	1,225	3.6	8,098

Source: Taken from table 1 of AIHW & AACR 1999.

Table 3.2: Classification of colorectal cancer primary site

Site	ICD-9 classification code	ICD-9 classification name
proximal colon	153.0	hepatic flexure
	153.1	transverse colon
	153.4	caecum
	153.5	appendix
	153.6	ascending colon
	153.7	splenic flexure
distal colon	153.2	descending colon
	153.3	sigmoid colon
rectum	154	rectum, rectosigmoid junction and anus

Note: Disease classified as 153.8 (other specified sites of the large intestine) or 153.9 (colon, unspecified) could not be assigned to either proximal or distal sites and so was excluded from all site specific analyses

Table 3.3: Comparison of TNM and Dukes' staging systems for colorectal cancer

Stage	TNM designation			Dukes' designation
0	Tis	N0	M0	–
I	T1	N0	M0	A
	T2	N0	M0	
II	T3	N0	M0	B
	T4	N0	M0	
III	Any T	N1	M0	C
	Any T	N2,N3	M0	
IV	Any T	Any N	M1	D

Notes: Tis, in situ; T1, tumour invades submucosa; T2, tumour invades muscularis propria; T3, tumour invades through muscularis propria; T4, tumour invades serosa, nodes, and adjacent organs; N0, negative lymph nodes; N1, 1–3 positive nodes; N2, >3 positive nodes; N3, positive nodes on vascular trunk; M0, no distant metastases; M1, distant metastases.

Source: Taken from Table 1 of Winawer et al. (1997).

Table 3.4: Risk factors for colorectal cancer

Average risk
Age 50 years and over and asymptomatic
Increased risk
Inflammatory bowel disease
Chronic ulcerative colitis
Chronic granulomatous colitis
Adenomatous polyposis
Familial polyposis
Gardner's syndrome
Turcot's syndrome
Oldfield's syndrome
Juvenile polyposis
HNPCC
Lynch I
Lynch II
Family history
Colorectal adenomas diagnosed under age 60 years
Colorectal cancer
Past history
Colorectal adenomas
Colorectal cancer
Breast, ovarian and uterine cancer

Source: Taken from Table 2 of Winawer et al. (1997).

Table 3.5: Mean sojourn times in each cancer stage

Stage	Mean sojourn time (years)
A	2.00
B	1.00
C	1.00
D	0.75

Source: Eddy (1980) for stages A, B and C—the value for stage D is an arbitrary assumption.

Table 3.6: Number of new cancers by year of diagnosis, primary site, sex, age at diagnosis and stage at diagnosis

Year	Site	Sex	Age	Stage A	B	C	D
1980–1989	Rectal	Male	LT 50	7	5	8	8
			50-59	9	18	20	14
			60-69	24	29	28	21
			70-79	25	35	32	20
			80 +	7	17	19	7
		Female	LT 50	4	7	6	3
			50-59	12	11	5	9
			60-69	19	18	17	9
			70-79	13	23	15	19
			80 +	7	13	10	12
	Distal	Male	LT 50	2	2	3	3
			50-59	6	6	5	14
			60-69	10	24	12	9
			70-79	10	24	14	25
			80 +	2	6	2	6
		Female	LT 50	0	2	5	2
			50-59	6	8	7	4
			60-69	13	24	16	10
			70-79	11	31	19	6
			80 +	5	19	14	9
	Proximal	Male	LT 50	1	8	7	3
			50-59	1	15	9	13
			60-69	3	21	20	17
			70-79	7	28	16	16
			80 +	5	17	15	7
		Female	LT 50	2	2	9	3
			50-59	1	7	20	11
			60-69	3	18	20	14
			70-79	5	37	36	24
			80 +	2	26	31	21

(continued)

Table 3.6: Number of new cancers by year of diagnosis, primary site, sex, age at diagnosis and stage at diagnosis *(continued)*

Year	Site	Sex	Age	Stage			
				A	B	C	D
1990–1996	Rectal	Male	LT 50	4	3	9	5
			50-59	5	12	15	13
			60-69	19	33	26	21
			70-79	19	26	29	17
			80 +	12	13	8	3
		Female	LT 50	2	3	7	4
			50-59	2	9	8	6
			60-69	23	18	24	16
			70-79	14	24	28	10
			80 +	11	17	18	10
	Distal	Male	LT 50	1	3	8	4
			50-59	1	10	9	8
			60-69	8	15	19	12
			70-79	13	30	15	17
			80 +	3	15	4	2
		Female	LT 50	5	7	3	1
			50-59	3	6	11	6
			60-69	9	13	12	9
			70-79	15	22	16	12
			80 +	7	12	15	6
	Proximal	Male	LT 50	1	9	5	4
			50-59	2	8	10	11
			60-69	5	25	12	19
			70-79	7	30	22	11
			80 +	3	20	8	8
		Female	LT 50	1	8	4	5
			50-59	5	10	10	5
			60-69	7	17	29	11
			70-79	9	34	34	17
			80 +	3	25	24	11
Total			Number	441	978	882	623
			Per cent	15.1	33.4	30.2	21.3

Source: South Australian hospital based cancer registry data

Table 3.7: Fitting stage at diagnosis, primary cancer site and age at diagnosis as ordinal variables

Model term	Deviance	Degrees of freedom	
Row effects model treating stage as ordinal	176.0	117	$p < 0.001$
Column effects model treating site as ordinal	73.3	59	$p = 0.100$
Column effects model treating age as ordinal	90.7	107	$p = 0.871$

Note: Each variable is fitted separately, so that the deviance represents a comparison with the fully saturated model

Table 3.8: Uniform association model of counts of new cases of colorectal cancer as a function of age at diagnosis, sex, year of diagnosis, primary site and stage at diagnosis

Model term	Deviance	Degrees of freedom	
(1) Saturated model	—	—	
(2) Uniform association model treating age and site as ordinal	132.0	131	$p = 0.458$
(3) Model (2) with all fourth & higher order interaction terms involving stage removed	147.6	146	
Change in deviance—(3)-(2)	15.6	15	$p = 0.409$
(4) Model (3) with site.sex.stage removed	158.1	149	
Change in deviance—(4)-(3)	10.5	3	$p = 0.015$
(5) Model (3) with age.sex.stage removed	158.1	149	
Change in deviance—(5)-(3)	10.5	3	$p = 0.015$
(6) Model (3) with all third order interaction terms involving year and stage removed	161.8	155	
Change in deviance—(6)-(3)	14.2	9	$p = 0.510$
(7) Model (6) with year.stage removed	165.1	158	
Change in deviance—(7)-(6)	3.3	3	$p = 0.345$

Table 3.9: Age adjusted estimates of clinical surfacing probabilities by primary cancer site, sex and stage at diagnosis

Site	Sex	Stage			
		A	B	C	D
Rectal	Male	0.20	0.30	0.29	0.21
	Female	0.21	0.27	0.31	0.20
Distal	Male	0.19	0.35	0.30	0.16
	Female	0.14	0.35	0.24	0.27
Proximal	Male	0.08	0.31	0.39	0.22
	Female	0.07	0.40	0.28	0.24
Total	Male	0.15	0.32	0.34	0.20
	Female	0.14	0.35	0.28	0.23
	Total	0.14	0.33	0.31	0.22

Table 3.10: Registrants with colorectal cancer and observed survivors, South Australian Hospital registry 1980–95

Stage	Year of registration	Proportion of total registrants who registered in this year	Number of observed survivors who registered in this year
A	1980	0.049	7
	1981	0.028	5
	1982	0.042	5
	1983	0.042	11
	1984	0.047	11
	1985	0.061	18
	1986	0.056	15
	1987	0.079	25
	1988	0.058	17
	1989	0.056	17
	1990	0.068	23
	1991	0.070	25
	1992	0.091	30
	1993	0.072	25
	1994	0.103	40
	1995	0.079	29
B	1980	0.054	11
	1981	0.047	13
	1982	0.063	23
	1983	0.052	11
	1984	0.054	21
	1985	0.042	17
	1986	0.054	27
	1987	0.052	26
	1988	0.053	29
	1989	0.060	26
	1990	0.058	33
	1991	0.076	46
	1992	0.100	67
	1993	0.083	60
	1994	0.087	67
	1995	0.067	58

Table 3.10: Registrants with colorectal cancer and observed survivors, South Australian Hospital registry 1980–95 *(continued)*

Stage	Year of registration	Proportion of total registrants who registered in this year	Number of observed survivors who registered in this year
C	1980	0.054	6
	1981	0.045	6
	1982	0.060	8
	1983	0.051	8
	1984	0.044	12
	1985	0.055	10
	1986	0.040	8
	1987	0.049	12
	1988	0.051	11
	1989	0.061	19
	1990	0.064	24
	1991	0.059	23
	1992	0.090	36
	1993	0.091	42
	1994	0.095	53
	1995	0.090	65
D	1980	0.035	0
	1981	0.078	1
	1982	0.038	0
	1983	0.037	0
	1984	0.042	2
	1985	0.053	2
	1986	0.069	5
	1987	0.048	1
	1988	0.072	0
	1989	0.081	3
	1990	0.069	4
	1991	0.069	2
	1992	0.080	2
	1993	0.077	6
	1994	0.077	12
	1995	0.077	25

Table 3.11: Number of observed deaths from colorectal cancer and from all other causes among registrants with colorectal cancer, South Australian Hospital registry 1980–95

Stage	Years on registry	Colorectal cancer deaths	Other deaths
A	1	11	16
	2	13	9
	3	11	5
	4	5	2
	5	3	3
	6	4	11
	7	1	7
	8	1	2
	9	3	8
	10	1	3
	11	0	1
	12	0	3
	13	1	1
	14	0	1
	15	0	0
	16	0	0
B	1	48	39
	2	56	19
	3	59	19
	4	33	12
	5	23	17
	6	9	11
	7	8	7
	8	3	4
	9	6	10
	10	2	7
	11	1	8
	12	2	4
	13	0	4
	14	1	3
	15	1	0
	16	0	0

Table 3.11: Number of observed deaths from colorectal cancer and from all other causes among registrants with colorectal cancer, South Australian Hospital registry 1980–95 *(continued)*

Stage	Years on registry	Colorectal cancer deaths	Other deaths
C	1	94	31
	2	147	18
	3	105	8
	4	37	7
	5	12	4
	6	13	5
	7	6	7
	8	5	4
	9	2	6
	10	3	3
	11	0	1
	12	0	1
	13	0	0
	14	2	1
	15	0	0
	16	0	0
D	1	254	14
	2	208	2
	3	50	3
	4	18	1
	5	4	1
	6	1	1
	7	2	0
	8	1	0
	9	0	0
	10	0	0
	11	1	0
	12	0	0
	13	0	0
	14	0	0
	15	0	0
	16	0	0

Table 3.12: Parameter estimates for the hypothesised survival distributions

Stage	Distribution and parameter	Estimated value
A	Exponential	
	λ	0.24
	Pareto	
	δ	4.00
	k	0.92
	Lognormal	
	μ	1.53
	σ	1.71
B	Exponential	
	λ	0.29
	Pareto	
	δ	19.76
	k	6.12
	Lognormal	
	μ	1.02
	σ	1.37
C	Exponential	
	λ	0.47
	Pareto	
	δ	19.46
	k	9.90
	Lognormal	
	μ	0.35
	σ	1.05
D	Exponential	
	λ	1.17
	Pareto	
	δ	2.71
	k	4.07
	Lognormal	
	μ	-0.67
	σ	1.07

Table 3.13: Summary of chi square goodness of fit tests comparing the number of deaths projected using the three hypothesised distributions with the observed number of deaths by stage

Distribution	Stage	Chi square value	Degrees of freedom	P value
Exponential	A	11.12	14	0.68
	B	16.02	14	0.31
	C	55.38	14	0.00
	D	593.33	14	0.00
Pareto	A	9.69	13	0.72
	B	15.63	13	0.27
	C	38.55	13	0.00
	D	12.13	13	0.52
Lognormal	A	9.48	13	0.74
	B	18.36	13	0.14
	C	25.68	13	0.02
	D	9.52	13	0.73

Table 3.14: Proportion lost to follow-up and cure rates by stage from lognormal model

Parameter	Estimated value (%)
Per cent lost to follow-up	3.9
Cure rate	
Stage A	64.0
Stage B	46.7
Stage C	25.9

Table 3.15: Estimated cure rates by stage from PROC LIFETEST

Stage	Estimated cure rate (%)	Cure rate adjusted for loss to follow-up (%)
A	79.3	78.5
B	58.8	57.1
C	32.7	30.0

Table 3.16: Polyp diameter at transition to invasive cancer

Polyp diameter on cancerous invasion (mm)	Number of polyps
10–20	2
20–30	3
30–40	7
40–50	5
50–75	3
total	20

Table 3.17: Results of fitting a Weibull distribution to polyp diameters on transition to invasive cancer

Parameter values	Diameter as size	Volume as size	Varying both λ and γ
λ	0.000588	0.000015	0.0000047
γ	2	3	3.3
chi square value	3.57	0.77	0.60
p value for goodness of fit test	0.31	0.86	0.74

Table 3.18: Results of model fitting for rates of polyp initiation and growth, by primary site and sex

3.18a: Males

Age group	0-39	40-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	
Proximal											
Frequency	83	189	211	301	468	490	532	474	323	184	
Fitted value	106	193	195	305	447	514	540	494	316	190	
χ^2 component	5.0	0.1	1.3	0.1	1.0	1.1	0.1	0.8	0.2	0.2	
Overall χ^2 value											9.8
Distal											
Frequency	59	186	210	369	490	572	596	484	270	115	
Fitted value	61	192	216	345	513	570	603	446	265	130	
χ^2 component	0.1	0.2	0.2	1.7	1.0	0.0	0.1	3.2	0.1	1.7	
Overall χ^2 value											8.3
Rectal											
Frequency	97	341	385	594	855	842	808	615	365	205	
Fitted value	92	324	395	608	855	852	827	644	364	187	
χ^2 component	0.3	0.9	0.3	0.3	0.0	0.1	0.4	1.3	0.0	1.7	
Overall χ^2 value											5.3

Table 3.18: Results of model fitting for rates of polyp initiation and growth, by primary site and sex*(continued)*

3.18b: Females

Age group	0-39	40-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	
Proximal											
Frequency	81	183	188	302	403	506	576	609	502	383	
Fitted value	84	175	184	270	385	514	589	606	484	411	
χ^2 component	0.1	0.4	0.1	3.8	0.8	0.1	0.3	0.0	0.7	1.9	
Overall χ^2 value											8.2
Distal											
Frequency	66	202	191	288	417	451	439	413	254	187	
Fitted value	60	195	184	289	451	467	446	380	238	185	
χ^2 component	0.6	0.3	0.3	0.0	2.6	0.5	0.1	2.9	1.1	0.0	
Overall χ^2 value											8.3
Rectal											
Frequency	106	244	236	311	452	512	558	490	336	329	
Fitted value	118	229	223	331	446	534	561	512	345	301	
χ^2 component	1.2	1.0	0.8	1.2	0.1	0.9	0.0	0.9	0.2	2.6	
Overall χ^2 value											9.0

Table 3.19: Parameter values for rates of polyp initiation and growth, by primary site and sex

3.19a: Males

Parameter	Site		
	Proximal	Distal	Rectal
Polyp initiation			
c_0	0.00295	0.00140	0.00215
c_1	-6.9	-8.2	-7.9
c_2	0.125	0.150	0.165
Polyp growth rate			
p	0.005	0.165	0.050
q	0.100	0.400	0.400

Table 3.19: Parameter values for rates of polyp initiation and growth, by primary site and sex *(continued)*

3.19b: Females

Parameter	Site		
	Proximal	Distal	Rectal
Polyp initiation			
c_0	0.00230	0.00065	0.01200
c_1	-7.0	-8.0	-7.1
c_2	0.120	0.165	0.135
Polyp growth rate			
p	0.025	0.160	0.020
q	0.200	0.200	0.050

1. Figures

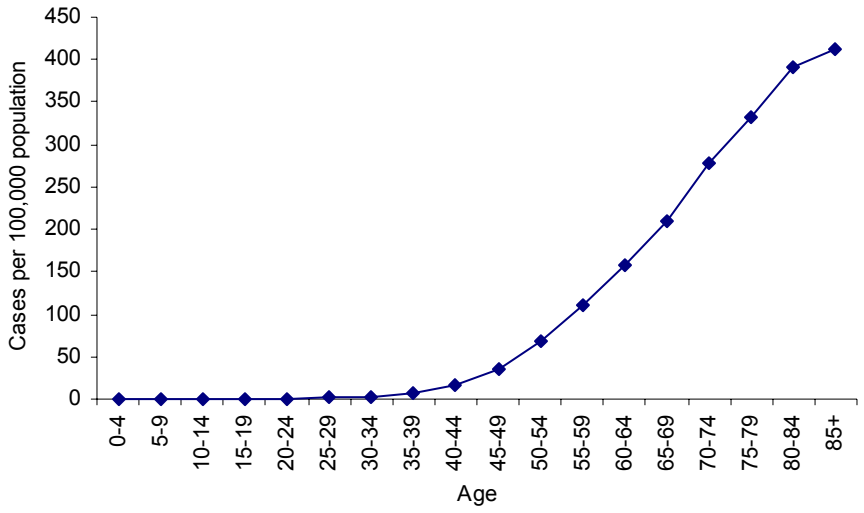

Figure 3.1: Colorectal cancer incidence by age, Australia, 1987–90

Source: National Cancer Statistics Clearing House

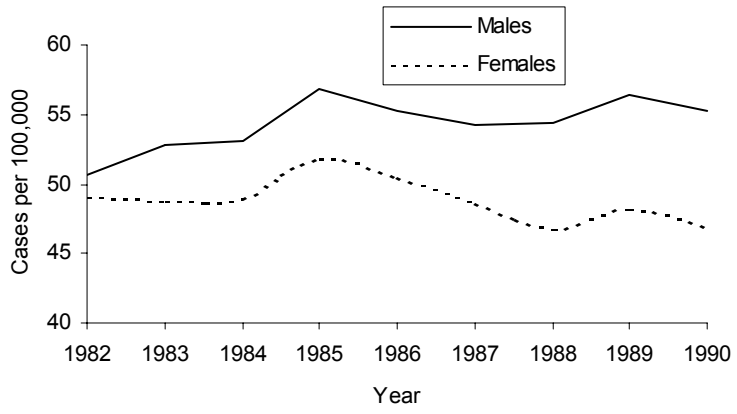

Figure 3.2: Colorectal cancer incidence in Australia, 1982 to 1990 by sex, age adjusted to the 1990 total Australian population

Source: National Cancer Statistics Clearing House

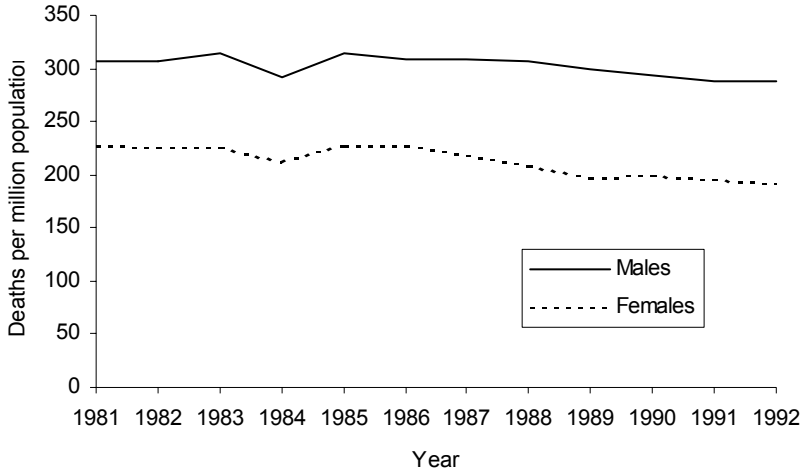

Figure 3.3: Colorectal cancer mortality in Australia, 1981 to 1992 by sex, age adjusted to the 1988 total Australian population

Source: Bennett et al. (1994)

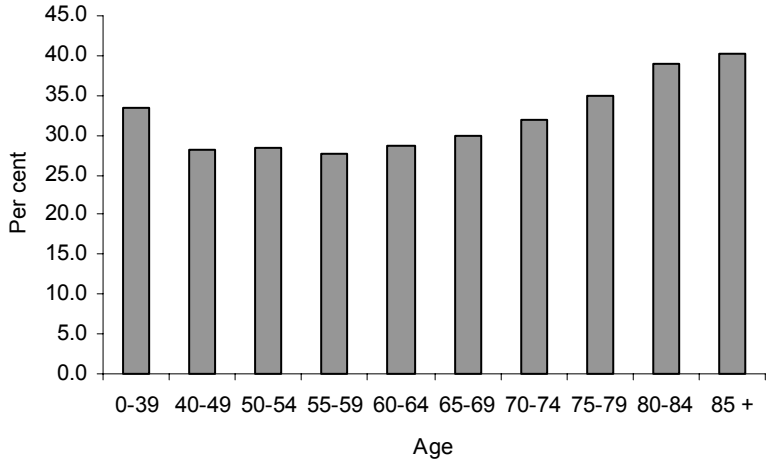

Figure 3.4: Proportion of colorectal cancers with primary site in the proximal colon by age, Australia, 1987–90

Source: National Cancer Statistics Clearing House

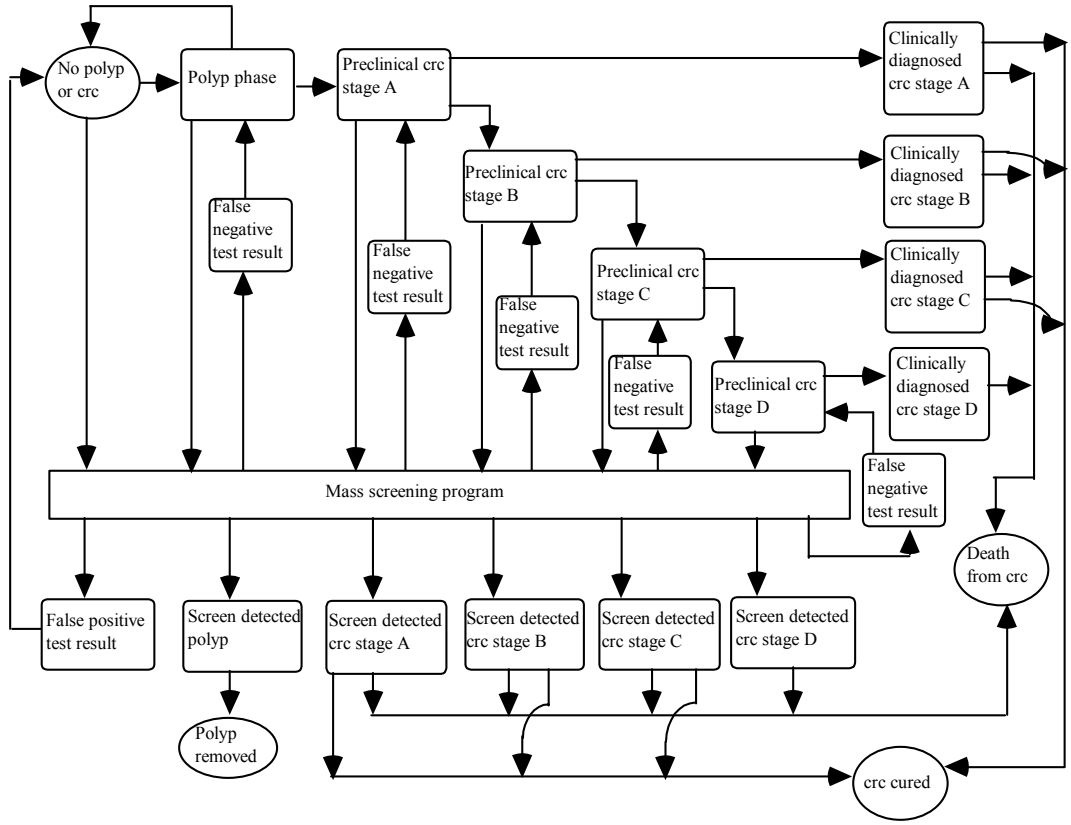

Figure 3.5: Structure of the disease and screening model for colorectal cancer

Notes:
1. Possible transitions between states are indicated by the arrows.
2. The state "death from other causes" has been omitted for clarity but may be reached from all other states.
3. This model formulation is similar in approach to that proposed for breast cancer by van Oortmarssen et al (1990).

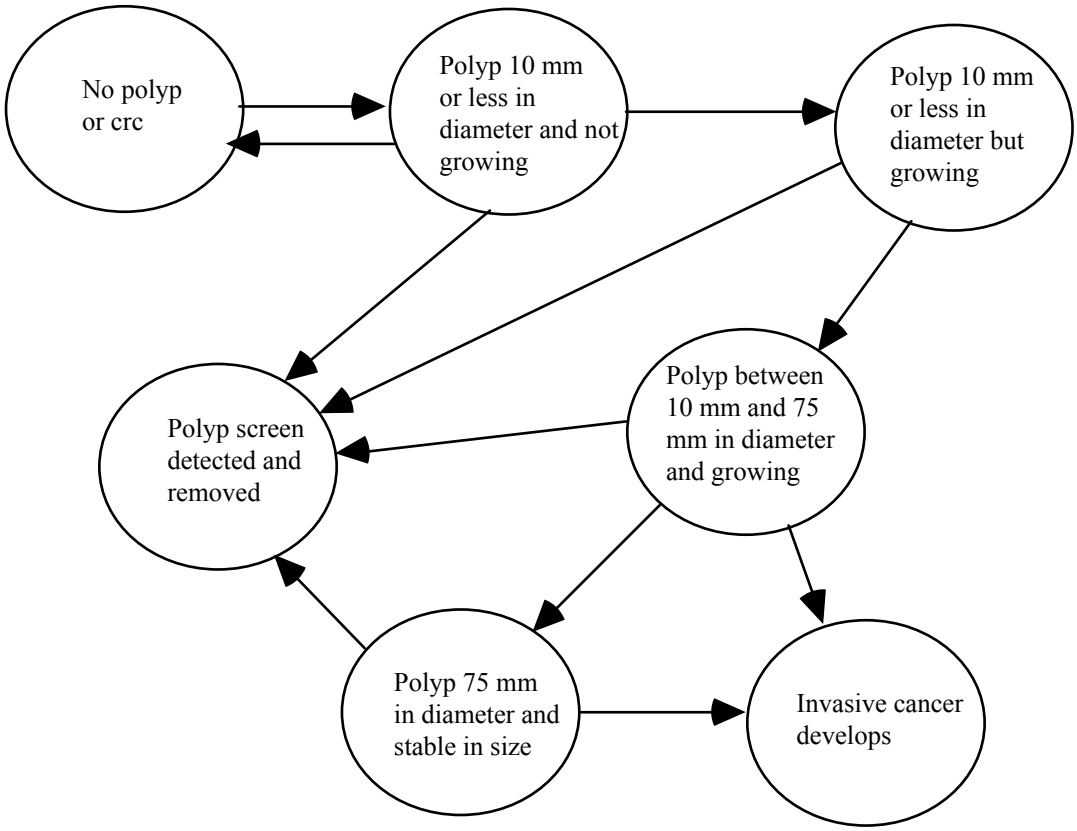

Figure 3.6: Detailed structure of the polyp phase of the disease model

Notes:
1. Possible transitions between states are indicated by the arrows.
2. The state "death from other causes" has been omitted for clarity but may be reached from all other states.

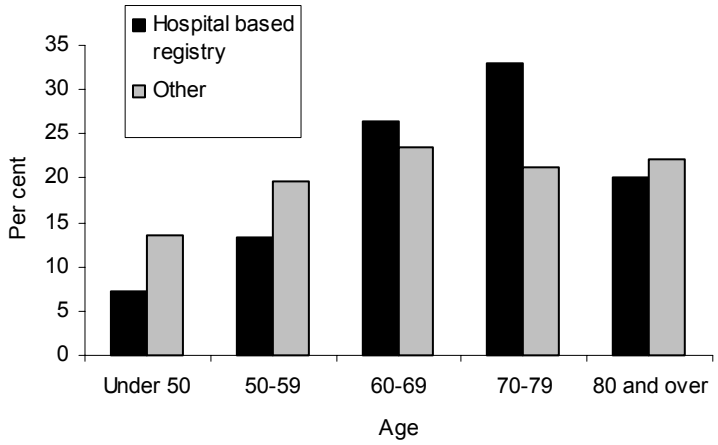

Figure 3.7: Per cent of cancers diagnosed in South Australia in the period 1982–1993 in each age group by whether or not the cancer was recorded in South Australian hospital based registries

Source: South Australian hospital based registries and National Cancer Statistics Clearing House

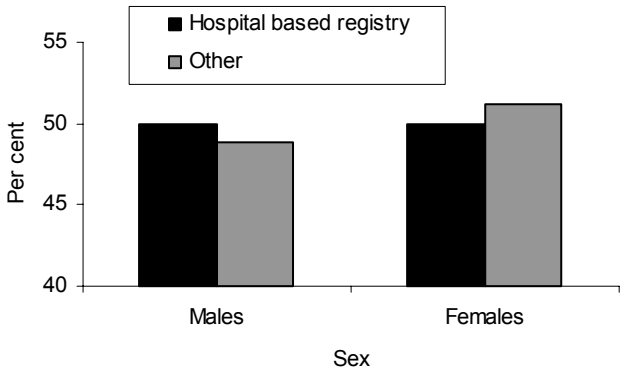

Figure 3.8: Per cent of cancers diagnosed in South Australia in the period 1982–1993 in each sex by whether or not the cancer was recorded in South Australian hospital based registries

Source: South Australian hospital based registries and National Cancer Statistics Clearing House

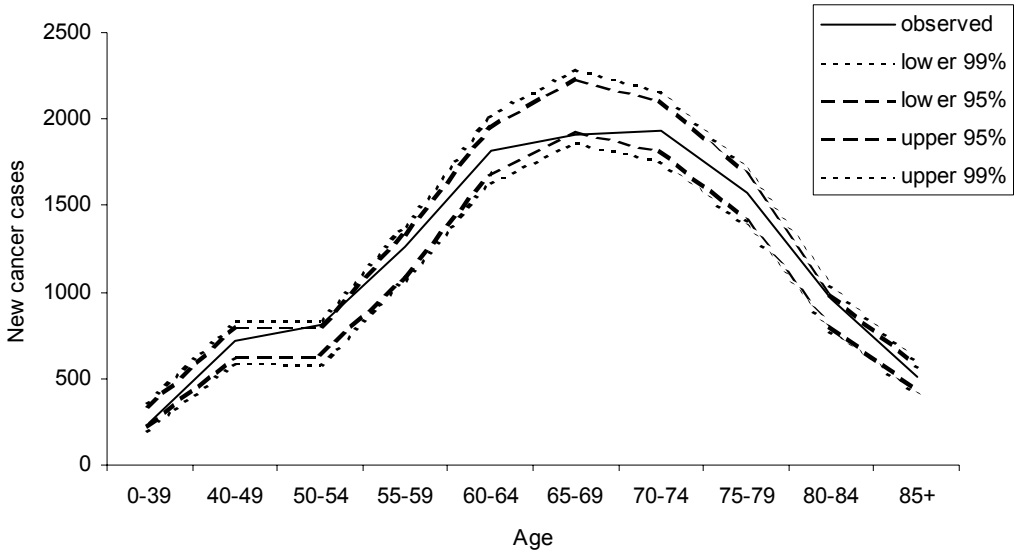

Figure 3.9: Observed counts of new cases of colorectal cancer and simulated confidence bounds, males, 1987–89

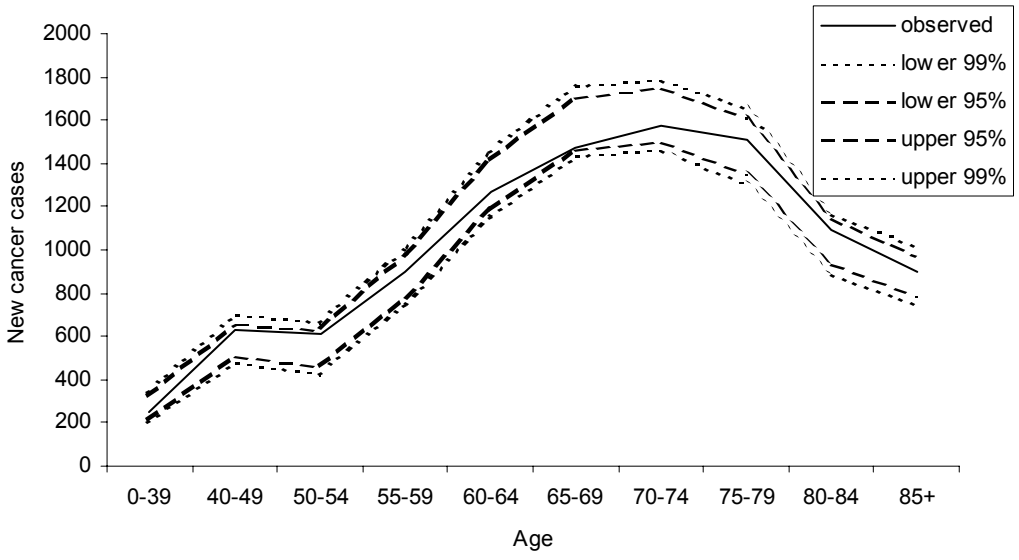

Figure 3.10: Observed counts of new cases of colorectal cancer and simulated confidence bounds, females, 1987–89

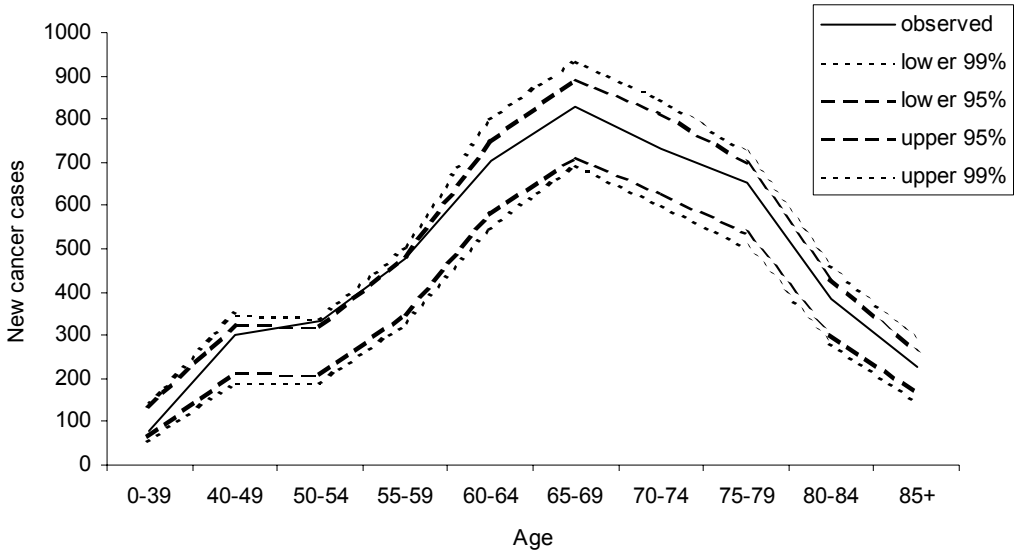

Figure 3.11: Observed counts of new cases of colorectal cancer and simulated confidence bounds, males, 1990

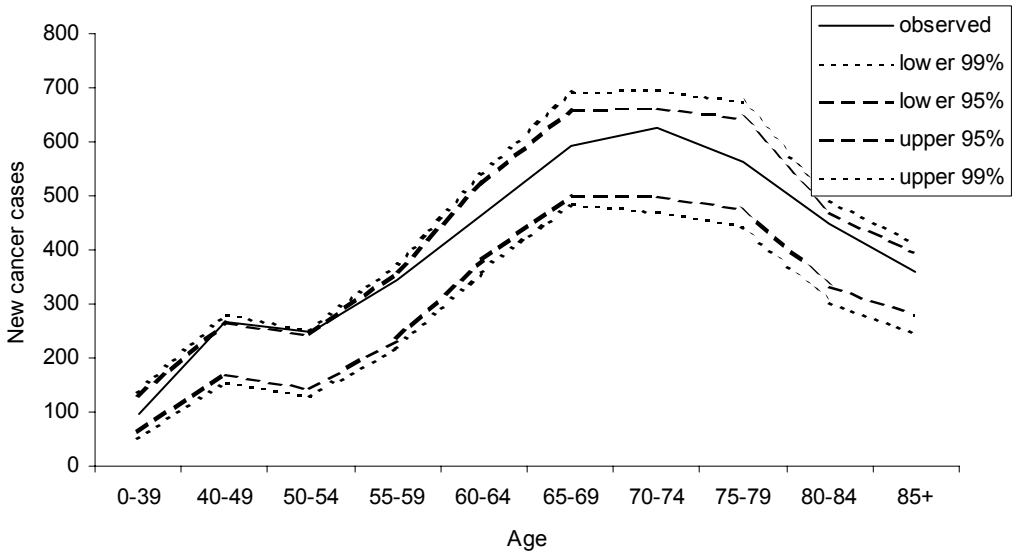

Figure 3.12: Observed counts of new cases of colorectal cancer and simulated confidence bounds, females, 1990

4 A model for screening for colorectal cancer

4.1 Screening tests

4.1.1 Faecal Occult Blood Test (FOBT)

Testing the faeces for occult blood has been done for more than 100 years, but credit for the recent interest in FOBT screening goes to Greegor (1967) who reported detecting a number of asymptomatic colorectal cancers using guaiac-impregnated cards (Knight et al. 1989). Although several different chemicals have been used to detect occult blood, the guaiac test Hemoccult II (SmithKline Diagnostics, Sunnyvale, California) continues to be the one most commonly recommended because of its high test specificity, simplicity and low cost—it is used in most large screening studies (St. John & Young et al. 1993).

The person using the test completes it by taking samples of stool from two different sites, using a wooden applicator stick, and smearing them thinly onto separate windows on the card incorporating the guaiac reagent. This is repeated for the next two bowel movements and the cards are returned to a screening centre, either in person or by mail (Winawer et al. 1997).

FOBT's face a number of problems with both sensitivity and specificity. Hemoccult II is based on a colour change which occurs due to the pseudoperoxidase activity of haemoglobin. It gives no indication of the amount of blood being lost. It is not specific for cancer since non-neoplastic lesions such as gum disease, gastritis, peptic ulcer disease, and haemorrhoids can also cause gastrointestinal bleeding. Nor is the test specific for blood per se, because other substances with peroxidase or pseudoperoxidase activity (red meat, bacteria and some fruits and vegetables) can cause false positive reactions if they are present in the stool (Winawer et al 1997). Because of the high false positive rate due to other sources of peroxidase or pseudoperoxidase activity, screening subjects using Hemoccult II are usually advised to follow a restricted diet prior to the test.

Antioxidants such as ascorbic acid can block pseudoperoxidase activity, causing false negative results. Positive reactions are also known to revert to negative in slides stored for more than a few days prior to laboratory testing (Knight et al. 1989). The sensitivity of guaiac based tests is increased if the test slide is rehydrated with a few drops of water before adding the hydrogen peroxide reagent but the trade-off for this increase is a reduction in specificity.

Immunochemical tests have been developed for detecting human haemoglobin in faeces which have the advantage of chemical specificity for human blood, thus

avoiding the false-positive results from other sources. However, these can produce a high false-positive rate because of the low concentration of faecal haemoglobin which they can detect leading some reports to describe them unsuitable for colorectal cancer screening (Knight et al. 1989). Other studies suggest that they can be used with an acceptable sensitivity and specificity (St. John & Young et al. 1993, Frommer et al. 1988).

HemoQuant (SmithKline Diagnostics, Sunnyvale, California) is a test based on conversion of heme to fluorescent porphyrins which is unaffected by dietary peroxidases, specimen storage and hydration. It also has the advantage of giving an indication of the amount of blood loss. However, it is much more expensive than the guaiac based tests (Knight et al. 1989) and has been reported to be unsuitable for screening for colorectal cancer (St. John & Young et al. 1993).

All tests of the faeces for occult blood share the disadvantage that colorectal cancers and polyps bleed intermittently and blood is distributed unevenly throughout the stool. However, the amount of bleeding increases with the size of polyp and stage of cancer (Winawer et al. 1997).

FOBT provides only an indication of the possible presence of cancers and large polyps. People with positive results must undergo some form of diagnostic evaluation, usually consisting of either double contrast barium enema, with or without flexible sigmoidoscopy, or colonoscopy (Winawer et al. 1997).

4.1.2 Sigmoidoscopy

Sigmoidoscopy involves visualisation of the rectum and sigmoid colon. Three types of sigmoidoscope have been used for screening—the rigid 25 cm scope and the flexible 35 cm and 60 cm scopes. Flexible sigmoidoscopes have largely replaced the rigid ones because they allow clearer visualisation of the mucosa, allow examination of more of the bowel and are more comfortable for the patient. The 60 cm flexible scope visualises more of the colon than the 35 cm scope without clinically important increases in discomfort or complication rates. Sigmoidoscopes allow for biopsy specimens to be taken but do not usually allow polyp removal (Winawer et al. 1997).

Four difficulties arise with sigmoidoscopy as a screening tool when compared with FOBT. The first is that while FOBT can detect lesions throughout the colon, sigmodioscopy only reaches part of the colon. The average depth of insertion of the rigid sigmoidoscope is 20 cm, which allows examination to just above the rectosigmoid junction. This gives it the potential to detect 25% to 30% of all colorectal cancers. The 60 cm flexible sigmoidoscope, when fully inserted, reaches the proximal end of the sigmoid colon or higher in 80% of examinations, thus potentially detecting 60% or more of colorectal cancers (Selby & Friedman 1989). The 35 cm flexible sigmoidoscope can visualise, on average, 50% to 75% of the sigmoid colon which allows it to reach 30% to 40% of lesions (Winawer et al. 1997).

The second difficulty is that while FOBT may be self administered, sigmoidoscopy requires trained health professionals for its administration. Thus it is a more

expensive screening tool and requires the availability of trained screening professionals. The procedure takes eight minutes on average, depending on the clinicians experience, but clinicians who are not specialised in lower endoscopy may take as long as 15 to 20 minutes with a 60 cm scope. Twenty four to 30 examinations under instruction is considered adequate training but even after that skill varies substantially among physicians. There is good evidence that non physician health professionals can be trained to use both 35 cm and 60 cm flexible scopes for screening with detection rates and complication rates equivalent to those achieved by medical endoscopists (Winawer et al. 1997).

The third difficulty is that of acceptability to the patient. Sigmoidoscopy requires preparation of the distal bowel by giving a saline laxative enema or similar preparation one to two hours before the procedure. Patients are not sedated and 10% to 15% experience at least moderate discomfort during the procedure. Reported rates of participation in screening sigmoidoscopy have been as low as 1.3% among people at average risk (Winawer et al. 1997).

The fourth difficulty is the relative safety of sigmoidoscopy. FOBT is a non-invasive procedure, so we would not expect adverse physical side effects (though such side effects may arise in whatever follow-up procedure is used with people with positive FOBT results). However there is a small possibility of adverse complications arising from sigmoidoscopy.

The major complication of sigmoidoscopy is perforation of the colon. Perforation rates of between one and two per 10,000 examinations have been reported, though there are no reported cases of consequent deaths (Winawer et al. 1997). Such perforations are obviously related to the experience of the practitioner, but studies with recently trained primary care physicians and nurse practitioners show no greater levels of complications (Selby & Friedman 1989).

Sigmoidoscopy has the advantage over FOBT that it is highly sensitive to both cancerous and precancerous lesions within its reach and, since it includes a biopsy as part of its operation, specificity is not an issue (Selby & Friedman 1989).

4.1.3 Colonoscopy

Colonoscopy, like flexible sigmoidoscopy, involves visualisation of the colon and rectum using a flexible endoscope. However unlike sigmoidoscopy it has the potential to visualise the entire colon. Further it is the only technique currently available which has the potential both to find and remove premalignant lesions throughout the colon and rectum (Winawer et al. 1997).

The performance of colonoscopy has mainly been studied in the context of diagnostic evaluation and surveillance rather than screening. In these cases the entire colon is reached in 80% to 95% of procedures, the depth of penetration depending mainly on the experience of the endoscopist and the adequacy of bowel preparation (Winawer et al. 1997). However, one study of screening colonoscopy showed that 98.6% of examinations reached the entire colon (Rex et al. 1991). Where

colonoscopy is incomplete, the addition of a double contrast barium enema is recommended to screen the entire colon (Fleischer et al. 1989).

Apart from its ability to reach the entire colon and its ability to remove premalignant lesions, colonoscopy has similar advantages and disadvantages to sigmoidsocopy as a screening tool. It also requires trained health professionals for administration, though unlike sigmoidoscopy it is currently performed only by physicians. It has similar problems of acceptability. The main complication is also perforation of the colon which is more common in colonoscopy than sigmoidoscopy. Mandel reports a rate of around three per 10,000 examinations in a screening programme trial (Mandel et al. 1993). However these examinations occurred in a major medical centre, where complication rates may be lower than in a general medical practice. Winawer et al. (1997) discuss the results of six prospective studies of colonoscopy which together suggest a perforation rate of one per 1,000 examinations.

Colonoscopy has the advantages of high sensitivity and the ability to remove premalignant lesions. As with sigmoidoscopy, it includes a biopsy as part of its operation so specificity is not an issue.

4.1.4 Barium Enema

Barium enema involves visualisation of the colon and rectum via radiography. It can be performed in two ways: as a single contrast study using barium alone, which reveals filling defects, or as a double contrast (DCBE) study, in which air is instilled after most of the barium has been removed and lesions in the mucosa are outlined by the retained barium. DCBE is slightly more difficult and expensive, but it is better at detecting mucosal lesions, including small polyps. Hence DCBE is usually preferred to single contrast studies as a test for cancer. DCBE can visualise the entire colon but studies show that 5% to 10% are unsatisfactory, requiring another attempt or a colonoscopy (Winawer et al. 1997).

The major complication of DCBE is also bowel perforation, but data on perforation rates are sparse. Patients are exposed to radiation during the examination, but considering the age and frequency at which such screening is usually recommended the consequent lifetime radiation dose would be lower than that for screening mammography. There is no direct evidence that DCBE, or other radiation of similar dose, frequency and anatomical coverage, causes clinically important increases in the risk of cancer or other tissue damage (Winawer et al. 1997).

Obviously DCBE does not include biopsy or removal of premalignant lesions. In other respects it is usually regarded as an alternative to colonoscopy. Winawer et al. argue that the two procedures have comparable sensitivity, cost-effectiveness and patient acceptability (Winawer et al. 1997). Further, the Australian Health Technology Advisory Committee suggest that colonoscopy has gradually replaced DCBE as the preferred examination of the large bowel since the mid 1980's with a concomitant decline in training, technical expertise and resources for the procedure (Australian Health Technology Advisory Committee 1997). Because of this I will

assume that colonoscopy is preferred to DCBE both as a follow-up to a positive FOBT test and as a primary screening tool.

4.1.5 Digital Rectal Examination

It has been common practice to screen for colorectal cancer by digital rectal examination, followed by tests for faecal occult blood if stool is present in the rectum. However, only 5% to 10% of all cancers are within the reach of the examining finger and stool samples obtained during the examination provide only a single specimen for faecal occult blood which may be falsely negative because of inadequate sampling (Winawer et al. 1997). I will not consider digital rectal examination as a potential screening methodology.

4.2 Overview of the screening model

4.2.1 Model description

The focus of this study is the modelling of an organised population screening programme. Figure 4.1 is a detailed diagram of the screening model. Each person is offered their first screen as part of the screening programme either (1) when the screening programme starts, if they are within the target age range at that time, or (2) when they reach the lowest age of the target range. They continue to be offered screening while they are in the target age range. If they refuse screening they continue to be offered screening annually. If they accept screening they continue to be offered screening after each screening interval.

People who are detected by screening with either a polyp or a cancer are allocated the prognosis associated with the cancer stage or polyp size by the natural history model described in the previous chapter. In principle, a cancer detected by screening in the same stage as it would have ultimately surfaced clinically may have an improved prognosis due to its earlier detection. There are no data on which to base an estimate of any such benefit, so the model assumes it to be negligible.

The focus of this study is on modelling the benefits of screening, but interpretation of these benefits would be incomplete without a consideration of screening costs. The cost estimates used here are taken from those presented by Salkeld et al. (1996). These cost estimates focus on the direct costs of providing screening. Costs associated with, for example, training health personnel to administer screening or publicity associated with persuading people to take part in screening are outside the scope of my model.

Both the costs and benefits of screening should be considered in the context of current screening practice. There is no current formal screening programme for colorectal cancer in Australia, but there is some use of screening among the general population. The costs and benefits of a screening programme are assessed relative to this current use of screening.

4.2.2 Modelling assumptions

This section presents a summary of the key modelling assumptions. A detailed discussion and justification for each assumption is given in the next section.

General assumptions

- Screening is only offered to people with no previous history of either colonic or rectal polyps or colorectal cancer;
- The model is governed by three screening probabilities, which may vary by age and sex but remain constant over time
  - the probability that a person offered screening for the first time will accept it and join the screening programme,
  - the probability that a person who refused the first screening offer will accept a subsequent offer and join the screening programme,
  - the probability that a person who accepted the first screening offer will accept a subsequent offer and remain in the screening programme.
- A person who leaves the screening programme has the same probability of rejoining it as someone of the same age and sex who never joined the programme;
- A person who joins the screening programme after previously refusing screening has the same probability of remaining in the programme as someone who joined it on the first offer of screening;
- Cancers detected at screening are assumed to have a prognosis and survival time no worse than they had on clinical surfacing in the absence of screening.

Background screening

- The rate of background screening is 1.8%
  - this applies to both average and high-risk groups;
- Participants in background screening are assumed to be screened annually from the age of 50 as recommended by the NHMRC guidelines (NHMRC 1999);
- The sensitivity and specificity of FOBT is assumed the same for background screening as for screening as part of a formal programme;
- An organised screening programme subsumes all background screening.

Costs of disease and screening

- The cost of detection of a screen detected cancer is assumed to be the cost of that screening programme;
- The cost of cancer treatment depends only on the stage of the cancer at detection—for a given stage it is the same irrespective of whether the cancer was screen detected or clinically surfaced.

4.2.3 Application of the model

The screening programme is assumed to operate from the base year (1990). The population is divided into single year age cohorts with age as at the base year. The model then calculates costs with and without a screening programme and the benefits of the screening programme. The costs are referred to as the *screening cost* and the *background cost* respectively.

The background cost is calculated first. Each member of a cohort is allocated randomly to background screening or no screening. Those allocated to background screening are assumed to have annual FOBT screening from the age of fifty or their age in the base year, whichever is higher, until they die or develop cancer. Each FOBT test may lead to a colonoscopy with a probability governed by the specificity of the FOBT test. If cohort members do not develop cancer, their contribution to the background cost is the sum of the costs of all their FOBT tests plus the cost of any colonoscopy. If they do develop cancer, their contribution to the background cost includes these costs plus the cost of treating their cancer. Those allocated to no screening contribute only the costs of any cancer treatment.

Screening operates for a specific time period from the base year, denoted the *screening period*, and is targeted at a specific age range. Hence the screening experience of each cohort member will depend on their age at the base year.

Those who are older than the target age range at the base year will be excluded from the screening programme. Their screening cost will be the same as their background cost, except that they are not subject to background screening.

Those who are in the target age range at the base year will be randomly allocated to the screening programme or to the non-attenders group based on the screening rate applying to their age and sex group. Thereafter their path through the screening model will be in accordance with the model described in Figure 4.1. At each screen the costs of administering that screen will be added to the screening cost. All cancers, whether screen detected or detected outside the programme, are allocated the cost of treatment appropriate to their stage at detection.

Members of the cohort cease screening when they reach an age outside the target range or when the screening period ceases or when they die or are detected with cancer. No further background screening occurs among members of this group who survive to an age outside the target range.

Age cohorts who are under the target age range are aged without any background screening from the base year until they reach the lowest target age. Thereafter they proceed as described above.

Cancers which are detected as part of the programme are randomly allocated prognosis and survival times appropriate to their stage at detection. Cancers not detected by screening have the same outcome as they had in the absence of the screening programme.

At the end of the modelling process, each cohort member who develops cancer will have an age at death in the absence of a screening programme and an age at death in its presence. The difference (if any) between these two ages is taken as the years

of life saved by screening. Each of these members will also be recorded as having died or been cured of cancer in both the presence and absence of a screening programme. These are used to count the cancer deaths prevented by screening.

The model also specifies a *study period*. This is the period which starts at the base year and specifies the period during which all population members are followed up. Cancers, deaths and costs incurred are only counted for the study period. Hence the only deaths which contribute to the measures of deaths prevented or years of life saved are those which occur within the study period. This study period may be set to a relatively short period to allow the simulation to mimic a RCT or it may be set to a very long period to allow an entire cohort to be followed through their remaining lifetime.

The model can be applied to a single age cohort, a group of age cohorts or the entire population. Hence a single age cohort or a group of age cohorts can be exposed to entire screening programmes, allowing assessment of the benefits of screening to participants. Alternatively the effect on the entire population of a screening programme commencing in the base year can be simulated, allowing the investigation of start-up effects.

4.3 Parameters for the screening model

This section describes the derivation of parameters for the screening model. These parameters are summarised in table 4.1.

4.3.1 Test sensitivity and specificity

FOBT

Mandel et al. (1993) and Kewenter et al. (1991) advocate rehydration of the slide by the addition of a drop of water to increase the sensitivity. However this leads to a decrease in test specificity and later trials conducted by Hardcastle et al. (1996) and Kronborg et al. (1996) used unrehydrated slides. I will follow these studies in using sensitivity and specificity estimates for unrehydrated slides, though the effect of rehydration will be investigated as part of the analysis of FOBT screening.

One possible source of screen sensitivity estimates are the randomised controlled trials of FOBT screening. Mendel et al. (1993), Robinson et al. (1995) and Kewenter et al. (1988) calculate sensitivity estimates for unrehydrated FOBT tests based on the number of cancers clinically surfacing in a 1 year, 2 year and 27 month period respectively following a negative screen. Their estimates of the sensitivity of FOBT for colorectal cancer at any stage were 81%, 53.6% and 28% respectively. The difficulty with this approach is that the cancer may have developed after the screen, leading to an under estimate of the sensitivity, or it may have been present at the time of the screen but not surfaced till after the specified period, leading to an over estimate of the sensitivity.

Gyrd-Hansen et al (1997)apply a form of the Walter and Day model (Day & Walter 1984) to the Danish RCT data (Kronborg et al. 1996) to jointly estimate the sojourn time in the preclinical cancer phase and the sensitivity of FOBT testing relative to colorectal cancer at any stage. Their estimates for mean sojourn time and test sensitivity are 2.1 years and 62.1% respectively. This result does not rely on the cancer surfacing in a specific period, but it is not sufficient for my study since it neglects the polyp phase and does not provide specific results for each cancer stage.

Another potential source of sensitivity estimates is published studies of FOBT testing followed by colonoscopy. Most of these focus on symptomatic patients and hence are likely to result in higher sensitivity estimates than would have been found among the asymptomatic general population who are the focus of my model. Further, most do not provide results by cancer stage. One exception to this is the study by Ahlquist et al. (1993).

Ahlquist et al. selected a group of 1217 patients aged at least 18 years undergoing routine structural surveillance evaluations following curative resection of a colorectal tumour. Stools were collected and tested prior to routine annual endoscopic or radiographic surveillance evaluations for the first 3 postoperative years. Since the endoscopic or radiographic evaluations were performed on all subjects independently of the stool testing, these served as criterion standards against which to directly assess test validity.

Subjects were instructed to refrain from red meat, nonsteroidal anti-inflammatory drugs, iron supplements and vitamin C for 5 days prior to and during stool collection. They then took a set of stool samples and returned them to the researchers for analysis. This process is similar to that which would be applied in a population screening programme, so the resulting sensitivity and specificity estimates should be applicable to population screening. However, it is possible that a small study such as this may have achieved better compliance with collection protocols such as the observance of dietary restrictions and better quality control in the handling and analysis of the stool specimens than would be achieved in mass population screening.

The stools were tested using both Hemoccult II and HemoQuant tests. The observed specificity of the Hemoccult tests was 95.3% and the overall sensitivity against all colorectal cancers was 26%. The specificity of the HemoQuant tests depends on the cutoff level of milligrams of haemoglobin per gram of stool chosen as defining a positive test result. When this was chosen so that the test specificity was also 95.3%, then HemoQuant also had an overall sensitivity of 26%.

Table 4.2 lists Ahlquist et al.'s sensitivity estimates for cancers in Dukes stage A or B and for polyps by the size of the polyp for both Hemoccult and HemoQuant. Ahlquist et al.'s data also suggest a sensitivity of about 50% for new primary tumours with stage greater than A or B. However, this is based on only 4 cancers so it can not be regarded as reliable. The meta-analysis conducted by Shimbo et al. (1994), which included studies of symptomatic patients and results obtained from RCT data, suggests a sensitivity of 58.8% for cancers in stage B or C. In the absence of any better data, this value will be taken as the sensitivity for cancers in stage C.

The sensitivity of FOBT tests for stage D cancer is arbitrarily set to 1. Since most of the benefit of screening comes from detecting early cancer (i.e. stage A or B), the effect of misspecification of the sensitivity for stages C and D should be small.

The use of Ahlquist et al.'s data in the model present some problems. The estimate of 6.3% for polyps under 5 mm in width is higher than the sensitivity estimate for polyps between 5 and 9 mm in width (5.6%), which seems an unlikely result. Further, the sensitivity estimates do not distinguish between stages A and B. Stage A cancers have higher cure rates and better survival than stage B but FOBT tests are likely to be less sensitive to cancers in stage A than B. The model parameters should reflect this.

One approach to these problems is to compare the model results with actual screening data. Rae has reported the results of a screening programme in north eastern New South Wales for the period 1987 to 1996 (Rae 1998). His results are not presented in enough detail to provide direct estimates of the model parameters, but he does present details of the stage distribution of cancers found by FOBT screening.

Rae's description of the screening programme is not detailed enough to simulate it directly. Instead a screening programme was simulated for a hypothetical population with the age and sex structure of the total Australian population but one tenth its total size. The simulation followed Rae's programme in applying screening to all people aged 40 years and over. It used the participation rates listed in table 4.1 and the sensitivity estimates given above for stage C and D cancers. The sensitivity rates for polyps and stage A and B cancers were varied systematically, using Ahlquist et al.'s rates as base values, until the stage distribution matched that reported by Rae. The results of this comparison are presented in table 4.3 and the modified estimates are presented in table 4.4. These are the sensitivity estimates used in the model. The effect of varying these parameters will be investigated in the sensitivity analyses.

Screening specificity may depend on participant compliance with a restricted diet. As noted above, some diets may produce false positive results and in a mass screening programme it is impossible to ensure that all screening participants follow the dietary regime correctly. However, Fleischer et al. claim that this is only important for rehydrated slides (Fleischer et al. 1989). They state that the average cooked diet has little effect on the results for unrehydrated slides. Similarly Norfleet demonstrated no statistically significant difference in specificity between FOBT tests applied to two groups of patients—one with and one without a restricted diet (Norfleet 1986). In view of this I have neglected the effect of dietary factors and taken the specificity to be that presented Ahlquist et al.'s results—95.3%.

The specificity estimate has no effect on the direct benefit estimates from the model. It does, however, effect estimates of both the cost and the possible harm from side effects of follow-up colonoscopy.

Ahlquist et al. note the possibility of different sensitivity to cancers and polyps in different parts of the bowel caused by the degradation of heme during colonic

transit. However, there are insufficient data to model such a variation in sensitivity, so I will neglect it for the purposes of my model.

Colonoscopy

Colonoscopy has the ability to screen the entire colon, so its sensitivity depends on the proportion of lesions missed in the examination. Hixson et al. prospectively examined 90 patients by tandem colonoscopy performed by two alternating examiners. A total of 15% of polyps under 10 mm were missed by the first examiner but no cancer and no polyp greater than or equal to 10mm in diameter was missed. So the model will use a sensitivity of 100% for cancers and polyps 10mm or bigger and 85% for polyps under 10mm.

Specificity is not an issue since a biopsy is available with colonoscopy.

Sigmoidoscopy

The sensitivity of sigmoidoscopy depends on both the proportion of the colon examined and the proportion of polyps or cancers missed in this part of the colon. My analysis of screening sigmoidoscopy will focus on the 60 cm flexible sigmoidoscope. As noted above, this reaches the proximal end of the sigmoid colon or higher in 80% of examinations. There are no data to model the proportion of the sigmoid colon missed in the remaining 20%. Instead I will assume that each examination visualises the whole sigmoid colon but no higher. Hence sigmoidoscopy will cover cancers in the rectum and part of the distal colon but not the proximal colon.

I have preclinical incidence estimates for the whole distal colon but not separately for the sigmoid colon. Table 4.5 presents the number and proportion of distal colon cancers clinically surfacing in the sigmoid colon and the descending colon (the remaining part of the distal colon) during the period 1982 to 1990. I will assume that these proportions also apply to polyps and use them as the proportion of distal colon cancers within the reach of sigmoidoscopy. I will assume that the sensitivity of sigmoidoscopy for these cancers is the same as that of colonoscopy—100% for cancers and polyps 10mm or bigger and 85% for polyps under 10mm.

As with colonoscopy, specificity is not an issue because a biopsy is available with sigmoidoscopy.

4.3.2 Screening participation rates

This section addresses participation rates for screening the whole population. Screening participation for population groups with different risk levels will be discussed in chapter seven.

FOBT

The best source of screening participation rates are the published screening study data, since these represent screening rates actually achieved in a mass screening programme. The disadvantage with these data is that willingness to participate in

screening is likely to vary with different countries and cultures and none of these estimates apply directly to Australia. However, there are no published participation data for FOBT based colorectal cancer screening programmes in Australia. The major published screening studies are:

The Minnesota RCT

This study recruited 46,551 participants age 50 to 80 years from volunteers from the American Cancer Society and fraternal, veterans and employee groups in Minnesota. They were randomised to screening once a year or screening every two years, using FOBT with colonoscopy follow-up of positive screens, or to a control group (Mandel et al. 1993).

The Nottingham RCT

This study recruited 152,850 participants aged 45 to 74 years who lived in the Nottingham area of the UK. Individuals were identified according to the general practice at which they were registered. Family Health Service Authority and general practice registers were used to compile a list of men and women in the target age range associated with each general practice. After removal of people judged unsuitable for the study (for example because of a previous diagnosis of colorectal cancer), the remaining individuals were randomised to a study group for screening using FOBT every two years or to a control group. Controls were identified but not told about the study, received no intervention and continued to use health-care facilities as usual. Screening group members were sent a Haemoccult FOB test kit, together with instructions and an explanatory letter from their family doctor inviting them to complete and return the test. For most of the study period people who did not accept the first invitation for screening were not reinvited for subsequent screens (Hardcastle et al. 1996).

The Funen RCT

This study recruited 140,000 people age 45 to 75 years living in Funen, Denmark. The study design was similar to that of the Nottingham study except that participants were selected from the population register of the county (Kronborg et al. 1996).

The Göteborg RCT

This study recruited 51,325 people aged 60 to 64 years from Göteborg, Sweden. Again the study design was similar to that of Funen and Nottingham except that all members of the screening group were invited for rescreening regardless of whether or not they participated in the first screen (Kewenter et al. 1991).

The Burgundy participation study

The objective of this study was to investigate the effect on compliance of demographic variables and of the way of proposing an FOBT in a colorectal cancer mass screening programme. Hence there was no control group and two recruitment

strategies were trialed—provision of the Haemoccult test kit by a medical practitioner during a consultation and direct mailing of the test kit.

The model requires three participation parameters—the proportion of the population who respond to the first screening invitation, the proportion of people in the screening programme who respond to second and subsequent screening invitations and the proportion of people who refused the first screening invitation who enter the screening programme at a later invitation.

Thomas et al. (1995) analyse longitudinal compliance with annual screening in the Minnesota RCT and Tazi et al. (1997) present a similar analysis for the Burgundy study of screening compliance. Both demonstrate that sex and age have a strong effect on participation rates, at least for the first invitation to screening. The Göteborg RCT focussed on people aged between 60 and 64 years of age which is too narrow an age range for use in this model (Kewenter et al. 1991). Of the others, the only one which reports participation jointly by age and sex is the Nottingham study (Hardcastle et al. 1996), so these are the participation rates listed in table 4.1 and used in the model. However, a comparison of the age specific participation rates of the four studies that report them, as given in table 4.6, is instructive. This shows that the Nottingham and Burgundy studies had similar participation rates, the Funen study had slightly higher rates and the Minnesota study had considerably higher rates.

The Minnesota RCT was based on a sample of volunteers from the American Cancer Society and fraternal, veterans and employee groups in Minnesota. Hence the participation rates would be expected to be higher than those for a mass population screening campaign. The Nottingham and Burgundy studies had similar compliance rates despite the trialing of different recruitment strategies in the latter study. The main difference in recruitment strategies between Funen and Nottingham is that participants in Nottingham received only one follow-up letter while those in Funen received two. The recruitment differences may have been influence by many things, including cultural differences between the two communities and different levels of awareness of the disease. I can't definitely ascribe the difference to this extra reminder letter, but it does suggest that measures such as this may have a significant effect on recruitment. Hence I will investigate the effect of increased participation in the sensitivity analyses.

Kronborg et al. (1996) and Tazi et al. (1997) present analyses of screening participation in second and subsequent rounds by participants in the first screen for the Funen and Burgundy studies respectively. These are reproduced in Tables 4.7 and 4.8. Table 4.7 suggests that between 93% and 95% of the people invited for a second or subsequent screen will accept the invitation. Further, this acceptance rate is relatively stable over the four screening rounds. Table 4.8 suggests that this value may be somewhat lower, but unlike table 4.7, these figures include people lost to follow-up as well as those refusing the screening offer. I will base my model parameter on Table 4.7 and assume that the probability of someone currently in the screening programme accepting a subsequent screening offer is 0.94.

For most of the Nottingham study only people who had accepted a screening offer were invited for subsequent screens. However, Hardcastle et al. report that later in the study screening invitations were sent to people who had refused an initial screen (Hardcastle et al. 1996). They report that 6.1% of people who refused an initial screen accepted a subsequent screening offer. I will use this value as the probability that an individual not currently in the screening programme will accept a subsequent screening offer.

It is likely that the proportions responding to second and subsequent screening offers will vary by age and sex, but none of the published studies present data which would allow the modelling of such variation. I will ignore it and assume that this proportion applies across all ages and for men and women.

Tazi et al. demonstrate a further variation in screening participation with screening interval, place of residence and source of screening invitation (Tazi et al. 1997). Again, there are insufficient data to directly model these variations, so I will ignore them.

Kronborg et al. report that 10.7% of the Funen study positive FOBT results were not followed up with a complete bowel examination. Mandel et al. report a corresponding 17.3% of the Minnesota study positive FOBT results not followed up by a complete bowel examination. Kronborg et al. further report that 2.8% of positive tests were not followed by any colorectal examination. Hardcastle et al., Kewenter et al. and Tazi et al. do not report the proportion of positive tests not followed up.

Neither Kronborg et al. nor Mandel et al. provide any details of the incomplete examinations, so it is impossible to model them directly. Instead I will use Kronborg et al.'s rate of 10.7% as the proportion who decline to follow up a positive test.

Sigmoidoscopy and colonoscopy

Very few studies have examined these screening modalities as part of a mass screening programme for people at average risk. Generally they are seen as either a follow-up to positive FOBT results or a screening modality for people at high risk. One exception to this is the study by Olynyk et al. (Olynyk et al. 1996). This study selected a random sample of men and women aged 55 to 59 years using a computerised database derived from the Western Australian Electoral Commission. Study subjects were recruited by mail with a follow-up telephone survey of non-responders to examine reasons for non-response.

Olynyk et al. estimated that there were 2881 eligible subjects. Only 11.9% (342) agreed to screening, but the non-responders survey suggested that with appropriate recruitment strategies this could be increased to 15.5%. The commonest reasons given for non-compliance were a lack of interest (30%) or a lack of time (28%). This suggests that non-compliance was not related to the particular screening instrument and hence would also apply to screening colonoscopy. I will take 15.5% as the first round acceptance rate for screening for each of these modalities.

A subsample of 77 study subjects filed in a brief acceptability survey. Of these, 76 said they would have the screening test again if required. This suggests that the

rates of second and subsequent screens for those already in a screening programme are similar to those for FOBT screening, so I will assume the same value (94%) for this rate. I will also arbitrarily assume the same rate (6.1%) for acceptance of second and subsequent screening offers by those not already in the screening programme.

4.3.3 Costs of screening and treatment

Screening

A detailed analysis of the costs of a screening programme is beyond the scope of this study. Instead I will focus on the direct health care costs of screening, diagnostic work-up, treatment, palliation and follow-up surveillance. Non health sector costs such as travel and personal expenses incurred by people being screened are not included, nor are indirect costs (such as productivity losses due to screening, diagnosis and treatment). Costs associated with the administration of a screening programme (such as the cost of mailing out test kits) are also not included.

Salkeld et al. estimated the components of these direct health care costs in Australia, and these estimates are presented in Table 4.9 after conversion from 1994 to 1990 costs using the total health price index published by the Australian Institute of Health and Welfare (1997). The relevant costs table in Salkeld et al. omitted a cost for barium enema, but their text described basing this cost on the Medicare Schedule fee so the cost figure for barium enema quoted in table 4.9 is taken from the Medicare fee schedule. I will assume that the only adverse side effect of colonoscopy and sigmoidoscopy is bowel perforation. A proportion of sigmoidoscopies and colonoscopies are randomly allocated a perforation using the side-effect rates presented above (2 per 10,000 for sigmoidoscopy and 1 per 1000 for colonoscopy). I will assume no deaths or permanent injury arise from this, so the only consequence is that the costs of treating a perforation are added to the screening costs.

As noted above, around 98.6% of screening colonoscopies succeed in visualising the whole colon. I will assume that in the remaining 1.4% of cases an additional DCBE was required and add the cost of DCBE to the screening cost.

Salkeld et al. used an estimate of 1.8% as the background screening rate (the proportion of people screened in the absence of a screening programme). They based this on American data as reported in Mandel et al. (1993) and in the absence of any Australian data I will also use this figure.

Treatment and follow-up

The cost of treatment depends on the stage of the cancer irrespective of whether or not it is screen-detected. Where screening detects a polyp I will assume a similar follow-up regime to Salkeld et al. Specifically, patients in whom large (greater than 10mm) polyps are found are assumed to have a colonosopy at three yearly intervals until they die from a cause other than colorectal cancer. This is assumed to prevent any further polyps or cancer so that the person remains cancer free over the rest of their life.

I will assume that successful treatment for cancer includes follow-up surveillance which prevents future cancers, so that patients cured of cancer are assumed to remain cancer free over the rest of their life.

Where a cancer leads to death, I will assume the additional cost listed in the table as advanced cancer.

Discounting of costs and benefits

In keeping with common practice in studies such as this, both the costs and benefits of screening will be discounted back to the base year. A wide range of discounting rates have been used in the cost-effectiveness literature. A common rate used in health cost-effectiveness studies has been 5% per year (see, for example, Wagner et al., 1996) but the US Panel for Cost-Effectiveness in Health and Medicine has recently suggested the rate should be 3% per year (Weinstein et al. 1996). They recommend that cost-effectiveness studies should use both 3% and 5% to keep new analyses comparable with past analyses. I will follow this recommendation and use 3% in the main analyses, with 5% as part of the sensitivity analyses.

4.4 Model validation and comparison with other studies

4.4.1 Model validation

In chapter one I noted that I would approach model validation using Eddy's proposed framework. That is:

- First order validation: which requires that the structure of the model make sense to people who have a good knowledge of the problem;
- Second order validation: which involves comparing estimates made by the model with the data which were used to fit the model;
- Third order validation: which involves comparing the predictions of the model with data which were available when the model was fitted but were not used in the estimation of model parameters;
- Fourth order validation: which involves comparing the outcomes of the model with observed data when applied to data generated and collected after the model was built (for example, data from a previously unobserved screening programme).

The extent to which each of the model parameters could be validated in this way varies. Many can only be regarded as having reached first order validation. For example, the mean sojourn times in each cancer state are based purely on expert opinion. Where data were available for parameter fitting, then second order validation has been applied. For example, the survival distributions for patients who die from cancer were tested against the data from the South Australian Hospital Registry that was used in their estimation. Third order validation was

applied to the disease incidence model as a whole, with 1990 incidence data which were not used in model estimation being compared to the model's projections. However, as Miller et al. (1990) note, an overall disease and screening model requires

"…validation with the best available data, which is preferably derived from randomised trials, before (it) could be extrapolated in ways that might guide policy decisions." (page 768)

As noted previously, there are four main published randomised controlled trials of FOBT screening for colorectal cancer. These are the Minnesota study (Mandel et al. 1993), the Funen study (Kronborg et al. 1996), the Newcastle study (Hardcastle et al. 1996) and the Göteborg study (Kewenter et al. 1991). This section presents a comparison of these studies' published results with simulations using my disease-screening model.

In each case, the disease-screening model was set up to simulate as closely as possible the screening protocol described in the published literature. The published screening participation rates were used and the model was applied to a simulated population with the published study population's age-sex distribution. Obviously this means that the screening participation rates used in my base model (listed in table 4.1) are not addressed by this validation. However, this can be regarded as a third order validation of the rest of the disease-screening model, since these published results were available at the time the model was constructed but not used in the parameter estimation other than for screening participation and follow-up rates. The specific parameters which should be regarded as excluded from third order validation are identified for each study below. No other RCT has published results since the model was constructed, so fourth order validation was not possible.

4.4.2 Mortality falls due to screening

The incidence of colorectal cancer varies between populations and, at least in the Minnesota study, the control group incidence varied from what would be expected from the general population. Hence an Australian based simulation model would not be expected to project the same numbers of cancers as found in RCT's based in other countries. However, if my model is accurately reproducing FOBT screening, then it should give approximately the same mortality reductions between the control and screened groups as found in the RCT results.

The Minnesota study was modelled as 10 years of screening over 13 years with annual FOBT testing and colonoscopy follow-up of positive results. There was a gap in the screening programme of three years following the fifth year of screening. In 5% of the cases, double contrast enemas were administered when colonoscopy was incomplete. This doesn't affect the simulated mortality fall, but does affect the simulated cost. Mandel et al. reported that 90.2% of people invited for screening accepted at least one screening offer, so this was used as the initial acceptance rate for both sexes at all ages. Acceptance of subsequent screens and acceptance of later

screens by initial refusers were modelled using the screening rates in table 4.1. Mandel et al. also reported 17.3% of positive FOBT results were not followed up with a complete bowel examination, so this figure was used instead of the rate in table 4.1. The study population was taken as the annual screening group listed in table 1 of Mandel et al. (1993).

The screen sensitivity and specificity values in table 4.1 are based on unrehydrated FOBT slides, but the majority of screens in the Minnesota trial used rehydrated screens. This was modelled by applying the ratio of the Minnesota study's sensitivity estimates for rehydrated and unrehydrated screening tests to the sensitivity parameters in table 4.1 and using the Minnesota study's specificity estimates directly in the simulation.

The Minnesota study also had a group undergoing biennial screening. However, this group started with much higher cancer rates in the screened group compared to the control group, which is a situation which my simulation model is not designed to replicate, so this other study was not included in the validation.

The Funen study was modelled as 10 years of biennial FOBT screening with colonoscopy follow-up of positive results. Kronborg et al. report 67% of people attending for the first round of screening, so this was used as the initial acceptance rate for both sexes at all ages. The acceptance rate for second and subsequent screens and the rate of positive FOBT results not followed up with a complete bowel examination, as listed in table 4.1, were derived from Kronborg et al.'s results, and so cannot be considered as part of a third order validation with this study. Only individuals who agreed to take part in the first screening round were invited for later screening, so the parameter for acceptance of later screens by initial refusers was set to zero. The study population was taken as the screening group listed in table 1 of Kronborg et al. (1996).

The Newcastle study was modelled as 14 years of biennial FOBT screening with colonoscopy follow-up of positive results. The study population was recruited throughout the first ten years of the trial, so that participants were offered screening between three and six times, depending on their date of entry into the study. Three of the trial group's publications described the study population at different times throughout the trial. These were used as a basis for modelling study participants' entry into the trial (Hardcastle et al. 1983, Hardcastle et al. 1989, Hardcastle et al. 1996). All the parameters listed in table 4.1 were applied to this simulation, but the initial screening participation rate and the subsequent screening rate for initial refusers were based on data published from this study. Hence they cannot be considered as part of the third order validation for this study.

The Göteborg study has not published results relating to differences in mortality between the screened and control group. Hence it was excluded from this comparison.

The model generated 199 simulations for each RCT and the proportional drop in mortality was recorded for each. The simulated results and the published result were combined into a set of 200 proportional mortality drops for each RCT. If the published result is consistent with the model, then it should lie within the

approximate 95% confidence limits represented by the fifth largest and the fifth smallest results in this set. Table 4.10 lists the published mortality fall for each RCT, along with the mean fall from the 199 simulations and this approximate 95% confidence interval.

All published falls are well within the simulated confidence interval for the model. Further, although the mean simulated result for the Minnesota study is slightly higher than the published result, the other two mean simulated results are very close to their corresponding published results.

All the model parameters except for the initial screening participation rates have been used in at least one simulated comparison with a trial which was not used in their estimation. Hence the model without the participation rates may be regarded as having passed a form of third order validation. If we include the participation rates, the model as a whole has passed a second order validation. However, none of these comparisons has addressed the issue of whether the participation rates would apply to screening in Australia. It is difficult to see how such as assessment could be done in the absence of a specific Australian mass screening programme. However, we can at least say that these participation rates represent rates which have been achieved in practice elsewhere and so should be possible to achieve in the Australian context.

4.4.3 Costs of screening

Validating the costs of screening is more problematic than validating the mortality benefits. There are a number of cost-effectiveness studies of FOBT screening for colorectal cancer leading to a variety of 'cost per life year saved' estimates. These estimates are mainly based on models of screening rather than specific RCT's and involve assumptions about costing screening services which may or may not apply to Australia.

The exception to this is the costs study by Salkeld et al. (1996), which used the results of the Minnesota annual FOBT screening study. Salkeld et al. estimated the costs of running this study if Australian costs had applied and the life years saved arising from the published mortality fall assuming Australian life expectancies.

In simulating the Minnesota study as described above, we also applied Salkeld et al.'s cost assumptions. The result of this was an estimate of a cost per life year saved which could be compared with Salkeld et al.'s estimate. The comparison between the simulated and published cost per life year saved is given in table 4.10.

This comparison says nothing about the validity of Salkeld et al.'s cost assumptions which we have used as the basis of my cost assumptions. However, Salkeld et al.'s method of estimating total costs and total life years saved was very different to my microsimulation approach. Hence the fact that his estimate and my simulated estimate are very close gives me confidence that the model's methods for accumulating the screening costs give reasonable answers. This can be regarded as a form of third order validation of the cost accumulation methods, though not of the

costs themselves. The costs listed in table 4.1 for use in my base model can only be regarded as having first order validation.

4.5 Tables

Table 4.1: Screening model parameters

Test sensitivity and specificity	
FOBT	**(%)**
Sensitivity	
Cancers by stage	
Stage A	20
Stage B	45
Stage C	59
Stage D	100
Polyps by size	
20 mm or greater	20
10–19 mm	11
5–9 mm	3
Under 5 mm	2
Specificity	
Screening programme	95
Background screening	95
Colonoscopy	**(%)**
Sensitivity	
Cancers and polyps 10 mm or greater	100
Polyps under 10 mm	85
Specificity	N/A
Sigmoidoscopy	**(%)**
Proportion of distal colon cancers within reach of sigmoidoscopy	
Males	86
Females	85
Sensitivity for cancers within reach of sigmoidoscopy	
Cancers and polyps 10 mm or greater	100
Polyps under 10 mm	85
Specificity	N/A

Table 4.1: Screening model parameters *(continued)*

Screening programme, average risk, recruitment rates	
FOBT	**(%)**
Proportion of people who accept first screening invitation	
Males	
45-49	34
50-54	51
55-59	52
60-64	54
65-69	54
70 and over	49
Females	
45-49	43
50-54	59
55-59	59
60-64	58
65-69	54
70 and over	47
Proportion of positive FOBT tests not followed up by a bowel examination	11
Background screening rate—all ages, both sexes	1.8
Sigmoidoscopy and colonoscopy	**(%)**
Proportion of people who accept first screening invitation, all ages and both sexes	15.5
All screening modalities	**(%)**
Proportion of people in the screening programme who accept second or subsequent screening invitation	94.0
Proportion of people outside the screening programme who accept second or subsequent screening invitation	6.1

Table 4.1: Screening model parameters *(continued)*

Screening and treatment costs	
Test or treatment	**Cost ($A 1990)**
FOBT	14
Colonoscopy	715
Sigmoidoscopy	88
Clinical examination	54
Perforation	10,828
Treatment (by stage)	
Stage A	13,324
Stage B	14,462
Stage C	14,631
Stage D	18,319
Advanced	21,300
Rate of bowel perforation for sigmoidoscopy	2 per 10,000
Rate of bowel perforation for colonoscopy	1 per 1,000
Discount rate for costs	3% per year

Table 4.2: Faecal blood test sensitivities for colorectal cancer and polyps from Ahlquist et al.

		Sensitivity(%)	
	Number of cancers or polyps	Hemoccult	HemoQuant
Cancers in stage A or B	14	29	29
Polyps by size			
20 mm or greater	10	20	20
10–19 mm	46	11	8.7
5–9 mm	107	5.6	5.6
Under 5 mm	223	6.3	2.2

Source: Taken from Table 1 of Ahlquist et al. (1993)

Table 4.3: Comparison of the stage distribution of screen detected cancers observed in the New South Wales study with that projected by the model

		Cancer stage			
	Polyp	Stage A	Stage B	Stage C	Stage D
Number of screen detected cancers generated by the model	8288	1144	867	484	169
Proportion observed in NSW study	75.7%	10.4%	7.9%	4.4%	1.5%
Expected number of screen detected cancers using NSW study proportions	8328	1088	875	498	163
Chi square values	0.20	2.87	0.06	0.41	0.24

Chi square test value	3.79
Degrees of freedom	4
p value	0.436

Table 4.4: Faecal blood test sensitivities for colorectal cancer and polyps modified using NSW study results

	Sensitivity(%)
Cancers by stage	
Stage A	20
Stage B	45
Stage C	59
Stage D	100
Polyps by size	
20 mm or greater	20
10–19 mm	11
5–9 mm	3
Under 5 mm	2

Table 4.5: The number and proportion of distal colon cancers clinically surfacing in the sigmoid colon and the descending colon by sex, 1982–90

	Males		Females	
	Number	Per cent	Number	Per cent
Descending colon	614	13.8	560	14.6
Sigmoid colon	3846	86.2	3282	85.4

Source: Australian Cancer Statistics Clearing House

Table 4.6: Percent screen compliance on the first screen by age at study entry for the Minnesota, Burgundy, Nottingham and Funen studies

Age	Minnesota	Burgundy	Nottingham	Funen
45–49		49.1		76.0
50–54	85.4	54.4	51.5	73.0
55–59	86.9	57.4	55.3	72.0
60–64	86.0	58.2	55.9	71.0
65–69	85.9	55.0	53.5	64.0
70 and over	79.1	46.5	47.2	50.0
All ages	**84.9**	**54.0**	**52.9**	**67.0**

Note: The rate for the Minnesota study for age 50 to 54 includes participants younger than age 50 years.

Source: Adapted from Table 4 of Thomas et al. (1995)

Table 4.7: Compliance during repeated screening from Kronborg et al.

Screening round	Number of people invited for screening	Number of people screened	Screening rate (%)
1	30762	20672	67%
2	20113	18781	93%
3	18236	17279	95%
4	16746	15845	95%
5	15279	14203	93%

Note: The screening rate is calculated as the number of people screened divided by the number of people invited for screening. This leads to slightly different numbers to those given in the original source.

Source: Table 2, Kronborg et al. (1996)

Table 4.8: Participation in successive screening campaigns after starting screening

	Participation (%) in			
Screening starts at	**2nd campaign**	**3rd campaign**	**4th campaign**	**5th campaign**
1st campaign	83.4	83.5	82.7	79.6
2nd campaign	-	87.2	85.6	82.5
3rd campaign	-	-	87.6	84.4
4th campaign	-	-	-	86.7

Source: Table 2, Tazi et al. (1997)

Table 4.9: Summary of costs of screening for and treating colorectal cancer

Test or treatment	**Cost ($A 1990)**
FOBT	14
Colonoscopy	715
Sigmoidoscopy	88
Clinical examination	54
Perforation	10,828
Treatment (by stage)	
Stage A	13,324
Stage B	14,462
Stage C	14,631
Stage D	18,319
Advanced	21,300

Source: Taken from Table 2, Salkeld et al. (1996)

Note: These figures have been converted from 1994 to 1990 costs using the total health price index published by the Australian Institute of Health and Welfare (1997).

Table 4.10: Comparison of simulated results with published results from randomised controlled trials

4.10.1: Mortality

Study	Published fall in colorectal cancer deaths (%)	Simulated fall in colorectal cancer deaths[a] (%)
Minnesota trial—Annual FOBT screening with colonoscopy follow up	32.2	35.6 (28.4–41.7)
Funen trial—Biennial FOBT screening colonoscopy follow up	17.7	18.7 (13.6–24.4)
Newcastle trial—Biennial FOBT screening colonoscopy follow up	14.3	14.8 (12.5–18.0)

4.10.2: Cost per life year saved

Study	Published cost per year of life saved ($A 1994)	Simulated cost per year of life saved ($A 1994)
Salkeld cost study based on Minnesota trial results	24,660	24, 055 (18,351–32,873)

Note: (a) The figures in brackets represent an approximate 95% confidence interval for the simulated result.

4.6 Figure

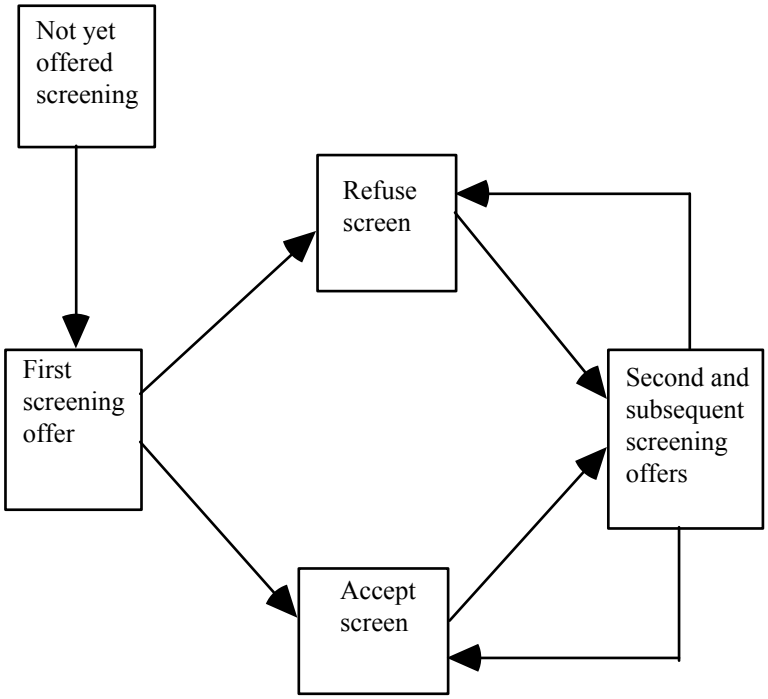

Figure 4.1: The screening model

Note: Possible transitions between states are indicated by the arrows.

5 Screening with FOBT

5.1 Introduction

This chapter will examine screening for colorectal cancer using the unrehydrated guaiac faecal occult blood test (FOBT) Hemoccult II with colonoscopy follow up of positive results. That is, I will assume that each participant is offered a single FOBT test, consisting of six faecal smears as described in chapter 4, and those with a positive result for any of the smears are offered a complete bowel examination with a colonoscopy. Cancers identified by screening are treated identically to cancers arising in the absence of screening. Polyps larger than 1 cm identified by screening are removed and the person followed up by colonoscopy at three yearly intervals. Polyps smaller than 1 cm are removed and the person goes back into the screening programme. The parameters governing the simulation will be those listed in table 4.1.

I will examine the costs and benefits to the general population of a mass screening programme, the benefits for individuals who take part in a screening programme and then the increase in benefits over time from the introduction of a screening programme. This analysis assumes screening by unrehydrated guaiac FOBT, so I will examine the effect of rehydrating the slides and the effect of using an immunochemical FOBT rather than the standard guaiac test. I will also examine the effect of combining sigmoidoscopy with FOBT screening as suggested by the NHMRC guidelines. Finally I will perform a sensitivity analysis on selected key modelling parameters.

5.2 Population based mass screening

5.2.1 Biennial screening

The age and screening protocols from the two population based mass screening RCT's were taken as the initial base model. That is, the model simulated offering people aged between 45 and 74 screens at two yearly intervals, though unlike the RCT's people were assumed to leave the screening programme when they reach 75. I compared the following protocols with this base model:

- biennial screening between the ages of 40 and 74;
- biennial screening between the ages of 50 and 74;
- biennial screening between the ages of 55 and 74;
- biennial screening between the ages of 45 and 69;
- biennial screening between the ages of 45 and 79;

- biennial screening between the ages of 45 and 84;
- biennial screening at all ages 45 and over; and
- annual screening between the ages of 45 and 74.

Table 5.1 lists the percentage fall in cumulative mortality and cumulative discounted life years lost and the cost per life year saved and per death averted for ten simulations of the base model applied to the whole Australian population. There is less than half a per cent difference between the highest and lowest values of the cumulative deaths prevented and YLS. The highest and lowest cost per life year saved differ by less than $200 while the highest and lowest cost per death prevented differ by less than $800. The differences between screening protocols in each of these measures would have to be much greater than this to have any practical implications.

Table 5.2 lists the simulated results for the base model and each of the alternate protocols. Figure 5.1 presents an analysis of the marginal costs and benefits of moving from the base model to each of the alternatives. The extra life years saved in moving to 40 years or lost in moving to 50 years as the lowest screening age are small and the cost per life year relatively large, suggesting that there may be little gain in biennial FOBT screening at ages below 50. However, the extra life years lost in moving from 50 to 55 is much larger and the cost per extra life year much smaller, suggesting that biennial FOBT screening should start at age 50.

There is a relatively large fall in life years saved with a small cost per extra life year in moving from 74 to 69 as the oldest screening age, suggesting biennial screening should continue at least to age 74. The extra years of life gained by moving the upper age limit to 79 is relatively small and this further decreases on moving the limit to 84. One would expect the gains in life years at these older age groups to be small but the cost per extra life year is also small and the fall in the number of deaths is relatively large, suggesting that screening may be worthwhile at these ages. However, the gains in both deaths prevented and life years saved in screening beyond 84 are small and the cost per life year saved is relatively large, suggesting that screening beyond 84 may not be worthwhile.

5.2.2 Annual screening

The final alternative protocol was a move to annual FOBT screening. This move generated a large saving in life years at a moderate cost per YLS. This suggests that annual screening may have significant advantages over biennial screening, so I repeated the above simulations with a base model of annual FOBT screening. In view of the small gains shown by biennial screening at ages below 50, I set the lower age limit for the base model at 50. However, I retained an upper age limit of 74 to facilitate comparisons with screening at older ages. This base model was compared with each of the following alternative protocols:

- annual screening between the ages of 40 and 74;
- annual screening between the ages of 45 and 74;

- annual screening between the ages of 55 and 74;
- annual screening between the ages of 50 and 69;
- annual screening between the ages of 50 and 79
- annual screening between the ages of 50 and 84; and
- annual screening at all ages 50 and over.

The results are presented in table 5.3 and figure 5.2.

There are moderate gains in life years saved in moving the lower screening age limit from 50 to 45 and 40 with a relatively high cost per extra life year, suggesting that the gains from annual screening below age 50 may not be worthwhile. However, there is a relatively large proportion of years of life lost by moving from 50 to 55 with a relatively small cost per extra life year, suggesting that screening should commence no later than age 50. As with biennial screening, there is a relatively large fall in life years saved with a small cost per extra life year in moving from 74 to 69 as the oldest screening age, suggesting annual screening should continue at least to age 74. Also there are moderate gains in life years saved and good gains in deaths prevented in moving the upper age limit to 84, with a cost per extra life year which is only a little above the average cost per life year for the base model. However, there are only small gains for a large extra cost in screening at ages over 84. This suggests that, as with biennial screening, the upper age limit for screening should be 84.

Thus I will use annual screening between ages 50 and 84 as my preferred model, with a simulated cumulative fall in deaths due to colorectal cancer of 34.2%, at a cost per death averted of $41,777, and a simulated fall in discounted YLL due to colorectal cancer of 28.5%, at a cost per YLS of $8,987. The proportion of screen detected cancers in each stage are presented in figure 5.3, along with the proportion of all cancers clinically detected in each stage in the absence of screening. This clearly shows the shift to earlier stages for screen detected cancers.

5.2.3 Comparison with other studies

It is interesting to compare my results with the results presented in table 4.10 for the overseas randomised controlled screening trials. My simulated fall in mortality is appreciably higher for biennial screening than either the Funen or Newcastle trials and appreciably lower for annual screening than the Minnesota trial.

The main reason for the difference for biennial screening is the length of the screening period. The Newcastle trial had 14 years of screening with participants recruited through the first ten years, so that some study participants had as little as four years of follow up. The Funen trial participants had 10 years of screening and follow up. Any study with a fixed screening and follow-up period will exclude mortality benefits from polyps and cancers detected during the study which would otherwise have led to deaths after the study period. The slow progression of colorectal cancer from polyp to invasive cancer means that this effect is larger for polyps, with most polyps taking between 10 and 15 years to progress to cancer

(Muto et al. 1975, Morson 1974). The relatively short study periods for the Funen and Newcastle studies mean that this had a larger effect on their estimates of mortality gain than for my simulation with its longer screening period.

The main reason for the difference for annual screening is the very high participation rate in the Minnesota study. The study participants for the Minnesota trial were recruited from volunteers from the American Cancer Society and from fraternal, veterans and employee groups. The self-selected nature of the group meant that participation rates were very high, with around 90% completing at least one test (Mandel et al. 1993). These are much higher rates than have been achieved in any of the population based mass screening programmes and they led to a higher proportion of the cancers being detected than in my simulated population screening programme, with a consequent increase in the mortality benefits.

There is also an appreciable difference in the cost per life year saved between my results and those of Salkeld et al (1996) based on the Minnesota study, as listed in table 4.10. As noted above, the Minnesota study had a higher participation rate which led to higher mortality gains and higher costs. However, the long term cost per life year saved should have been similar to ours. The difference arises from differences in the screening period. The Minnesota study had 10 years of screening over a 13 year period (Mandel et al. 1993). As noted above, the mortality gains from many of the screen detected polyps would not be apparent during this period, but the related screening costs would be counted. Hence the shorter period will lead to a higher cost per life year saved ratio.

5.3 The benefits of FOBT screening for screening participants

The focus of this analysis is the reduction in risk of death from colorectal cancer due to annual screening between the ages of 50 and 84. The average benefits of screening to screening participants were calculated by simulating screening for a synthetic cohort of one million men and one million women at ages 50 and 70 at the start of the screening programme with follow-up over the rest of each population member's life. I also simulated the percentage fall due to screening in cumulative colorectal cancer mortality and YLL for ten simulations of synthetic cohort aged 50 years at the start of screening. The range of these estimates gives an approximate idea of the variability in the simulation results due to the stochastic nature of the microsimulation approach. Table 5.4 lists the results. The simulated values for falls in mortality and YLL vary by 0.8% across the ten simulations.

Table 5.5 lists the simulated falls due to screening in both colorectal cancer deaths and YLL for screening starting at each of the selected ages. Obviously the larger gains come from starting screening at age 50. The average risk of death from colorectal cancer for screening participants over the rest of their life fell by 68%, while the average years of life lost due to colorectal cancer fell by 62%. However, the gains from screening remain substantial even for those who enter the screening

programme for the first time at age 70. For these people the average risk of death from colorectal cancer over the rest of their life fell by 52%, while the average years of life lost due to colorectal cancer fell by 45%.

This analysis is based on the risk of death from colorectal cancer over each participant's remaining lifetime. Obviously each person will die eventually, so the decrease in the risk of death from colorectal cancer represents a shift to other causes of death rather than a decrease in the overall risk of death. An equally important question is how much extra length of life was gained from this shift. Table 5.6 lists the average years of life gained for each true positive screen—where a true positive screen is taken as one which detected a cancer or a precancerous polyp (i.e. excluding screens which detected polyps which would not otherwise have progressed to a cancer). This includes longer survival for those who eventually die from cancer as well as extra years of life due to the prevention of death from cancer. Each positive screen resulted in an average of 2.1 extra years of life for people starting screening at age 50 and an extra 1.7 years of life for people starting screening at age 70.

5.4 The benefits of FOBT screening over time

I simulated the fall in both annual and cumulative mortality in the years following the beginning of a population screening programme using my preferred annual screening protocol from section 5.2. Figure 5.4 shows the fall due to screening in cumulative mortality from colorectal cancer by year from the introduction of screening and figure 5.5 shows the corresponding fall due to screening in annual mortality from colorectal cancer. Both these figures cover a period of 40 years from the introduction of screening.

The cumulative mortality gain rises sharply for the first 14 years to a value of around 24% and then continues to rise more slowly across the remaining 28 years till it reaches the value of 34.2% in the fortieth year.

The annual mortality gain, which is the difference in any specific year between deaths from colorectal cancer with and without screening, also rises sharply over the first 14 years to a value of about 35%. It then continues to rise more slowly to year 30. After year 30 it fluctuates around a value of about 39%.

I fitted a poisson regression model to the age-sex standardised colorectal cancer mortality rates from 1978 to 1989 and projected this forward to 1997. I then applied the simulated annual mortality falls from my model to the actual mortality rates for the period 1990 to 1997 and compared these with a one sided 95% projection interval from the regression. The results are summarised in figure 5.6. These suggest that the mortality rate with screening would become statistically significantly different from the projected rate during the third year of screening.

This means that a national monitoring system should be able to demonstrate a statistically significant fall in colorectal cancer mortality after three years of a

screening programme. However, the data in figure 5.6 demonstrate one difficulty in using mortality trends to monitor a screening programme. The observed mortality (in the absence of screening) is consistently lower than the projected mortality in the years following 1990 (the notional start year for the programme). This suggests that some other combination of factors was acting to reduce mortality in those years. Hence comparing the projected mortality with the simulated mortality in the presence of a screening programme in this case would overstate the effect of screening and lead to an earlier apparent identification of screening benefits.

Part of the purpose of monitoring would be to see if the screening programme is achieving its potential gains. The simulation suggests that at least 12 to 14 years of screening are needed before the actual mortality gains come close to the potential gains and a total of around 30 years before the screening programme actually achieves its full potential gains.

5.5 Suitability of FOBT screening for government funding

Table 5.7 lists the estimated cost-effectiveness of a number of health interventions which received government funding, adjusted to 1990 prices (AHMAC 1991). These estimates are based on a discount rate of 5% per year, so the estimated cost per life year saved from my model has been recalculated with a 5% discount rate for inclusion in the table. My estimated cost per life year saved falls well within the range presented in the table. Further it is within the range estimated for breast cancer screening and substantially lower than the cost per life year estimated for cervical cancer screening.

Table 5.8 lists the cost per extra life year saved and per death averted for annual FOBT screening from age 50 to 84 with colonoscopy follow up of positive screens and the marginal costs from changing the start and end screening ages and from moving from annual to biennial screening. Annual FOBT screening has an average cost per life year saved of $8,987 and cost per death averted of $41,777, both of which are well below the values found by George et al. (1998). The life years saved by moving the starting age for screening to 45 cost $31,751 per extra life year, which is very close to the limit of $32,000. Further the cost per extra death averted is $494,885, which is well above the limit of $260,000. The life years saved by moving the starting age from 55 to 50 cost $15,526 per extra life year. The life years saved by moving the finishing age for screening from 84 to death cost $51,111 per extra life year, which is well above the $32,000 limit, while the life years saved by moving the finishing age from 79 to 84 cost $13,630 per extra life year. The life years saved by moving from biennial to annual screening cost $16,721 per extra life year saved. These data suggest that, at least on cost-effectiveness grounds, my preferred annual FOBT screening programme would be a suitable health intervention for government funding.

5.6 Other faecal occult blood tests

5.6.1 Rehydrated guaiac FOBT

I have followed the practice of the more recent RCTs in modelling the use of unrehydrated slides in the FOBT. However, some authorities argue for the use of rehydration to increase FOBT sensitivity, though at the cost of a decrease in specificity (Mandel et al. 1993, Kewenter et al. 1991). This section examines a simulation of my preferred annual FOBT screening programme using rehydrated slides.

I will assume that the cost of a screen is identical for rehydrated and unrehydrated slides, so the only extra parameters I need to have are the sensitivity and specificity for rehydrated slides. The Minnesota study used both rehydrated and unrehydrated slides and presents sensitivity and specificity estimates for each (Mandel et al. 1993). As noted in chapter 4, the sensitivity estimates are based on the number of cancers clinically surfacing in a 1 year period following a negative screen. Hence while estimates for both rehydrated and unrehydrated FOBTs are likely to be high, the ratio between the two estimates may be more reliable. I derived a sensitivity estimate for rehydrated FOBT by multiplying my sensitivity estimate for unrehydrated FOBT by the ratio between Mandel et al.'s sensitivity estimates for rehydrated and unrehydrated tests. I also derived a specificity estimate for rehydrated slides from Mandel et al.'s specificity estimates in an analogous way. The results are given in table 5.9.

Table 5.10 lists the simulated results for both rehydrated and unrehydrated FOBT screening. The cost per YLS and per death averted for rehydrated slides is nearly twice that for unrehydrated slides. Further, the cost per extra YLS in moving from unrehydrated to rehydrated slides is $118,900, which is well above my $32,000 limit, with only an extra 2% YLS. These results support the use of unrehydrated rather than rehydrated slides in a population screening programme.

5.6.2 Immunochemical FOBT

Immunochemical FOBTs have been advocated as providing a better sensitivity and specificity than guaiac FOBTs (Frommer et al. 1988), though at a greater cost. St. John & Young et al. (1993) present a comparison of their sensitivity and specificity with the standard guaiac FOBT and Shimbo et al. (1994) include them in their modelling of cost-effectiveness of FOBT screening in Japan. This section examines a simulation of my preferred annual FOBT screening programme using an immunochemical FOBT.

The extra parameters I need for modelling immunochemical FOBT screening are the sensitivity, specificity and cost. Unfortunately there are no studies of population mass screening using immunochemical FOBT, so I must draw my estimates from published results of smaller studies.

Shimbo et al. give cost estimates for both standard guaiac and immunochemical FOBTs as part of their simulation modelling. I can't apply their cost estimates directly to my model since they are based on Japanese social insurance reimbursement tables, and hence may differ from Australian costs. However, their estimates put the cost of immunochemical FOBTs at three times the cost of standard guaiac FOBTs. So I can get an approximate cost of immunochemical tests by multiplying my standard FOBT cost by three. I will use this as my FOBT cost and assume all other costs remain the same.

Shimbo et al.'s sensitivity estimates for polyps do not differentiate by polyp size and so are unsuitable for use in my model. St. John & Young et al. present estimates of sensitivity for polyps less than 1 cm and greater than 1 cm. As noted in chapter 4, these estimates are likely to be too high because the study used symptomatic patients. However, I can apply the ratio of their estimates of sensitivity for immunochemical and standard guaiac FOBTs to my sensitivity estimates for standard guaiac FOBTs to give estimates of sensitivity for immunochemical FOBT for polyps. Their cancer sensitivity estimates do not differentiate by cancer stage but Shimbo et al. derive sensitivity estimates for cancers at stages A, B and C which I will use. As with guaiac FOBT testing, I will assume a sensitivity of 1 for stage D cancers. Shimbo et al. provide a specificity estimate for immunochemical FOBT of 99.1% based on an unpublished Japanese study. However, they do not provide any details of this study so I have no way of assessing its reliability. I will use instead St. John & Young et al.'s directly measured specificity estimate of 97.8%.

Table 5.12 lists the simulated results for both immunochemical and standard guaiac FOBT screening. The immunochemical FOBT has a cost per YLS which is around $1,000 lower than that for the standard FOBT. Further, it produces a fall in YLL which is around 10% greater. These results can only be regarded as indicative since the parameter estimates are only approximate and there are no studies of population based screening using immunochemical FOBT against which to validate the model. However, they do suggest that immunochemical FOBT should be seriously considered as part of population based screening for colorectal cancer.

5.7 Combining FOBT and sigmoidoscopy

5.7.1 Costs and benefits of population screening

Both Winawer et al. (1997) and the NHMRC guidelines suggest combining FOBT and sigmoidoscopy screening. They suggest that people be offered annual FOBT screening with sigmoidoscopy every five years. However, as Winawer et al. note,

'The individual components (of this strategy) are supported by strong evidence but the added value of combining the two, while theoretically present, is not well established by research evidence.' (p 598)

There have been no major published studies of the use of this combination in population screening. Hence, although the characteristics of each test are well

understood, there are no data on how well they operate in combination. In particular, although I have good data on screening participation rates for FOBT screening and some data on screening participation rates for sigmoidoscopy, I have no data on how acceptable screening participants would find the combination.

It is unlikely that participation in such a screening programme would be higher than for standard FOBT screening. So I could regard a FOBT screening programme where all participants underwent 5 yearly sigmoidoscopy as providing an upper bound to screening participation. Similarly, participants in a FOBT screening programme are not less likely to participate in sigmoidoscopy screening than the average population. Hence my participation estimate for sigmoidoscopy as the primary screening tool, listed in table 4.1, could be regarded as providing a lower bound.

I simulated two screening programmes with this combination. In both I assumed the same target age range as for my preferred screening programme and the same overall participation rates as for standard FOBT screening. All participants were offered a sigmoidoscopy screen on joining the programme and at 5 yearly intervals thereafter. People who left the programme and later rejoined were offered sigmoidoscopy on rejoining and thereafter at 5 yearly intervals. All participants were offered annual FOBT screens in the years between the sigmoidoscopy screens.

In the first simulation, all screening participants accepted the sigmoidoscopy. In the second simulation I assumed a sigmoidoscopy screening participation of 15% (the participation rate from table 4.1). Those who rejected sigmoidoscopy underwent FOBT screening instead. For example, the overall screening participation rate for males aged 50 to 54 is 51%. I assumed that, on joining the programme, 15% had a sigmoidoscopy while the remaining 36% had a FOBT. After five years, those who accepted the initial offer of a sigmoidoscopy and who were still in the programme accepted the second offer of a sigmoidoscopy with a probability of 0.94 (the probability of accepting second and subsequent screens from table 4.1). Those who rejected the initial offer of a sigmoidoscopy accepted the second offer with a probability of 0.061 (taken again from table 4.1). The screening programme proceeded in this way, with some participants accepting the 5 yearly sigmoidoscopy and the remainder opting for FOBT in the fifth year.

Table 5.13 presents the results of these simulations, along with the results from my standard simulation without sigmoidoscopy for comparison. If all screening participants accept sigmoidoscopy, the cost per YLS is around $2,000 lower than for FOBT screening without sigmoidoscopy. If sigmoidoscopy participation rates only reach 15%, the cost per YLS is virtually unchanged. However, the extra YLS ranges from around 2.6% for the lower sigmoidoscopy participation rate to almost 11% for the higher sigmoidoscopy participation rate.

These results suggest that the inclusion of an offer of 5 yearly sigmoidoscopy in an annual FOBT screening programme would result in a cost-effective improvement in colorectal cancer mortality.

5.7.2 The benefits to screening participants

This combination of modalities has not been well studied even in the absence of an organised screening programme. Hence it is worth asking how much extra benefit would screening participants receive if they follow the recommendations and add five yearly sigmoidoscopy to an annual programme of FOBT screening. I am on stronger ground in simulating such screening for screening participants because the individual effects of screening with each modality have been well studied and I no longer need concern ourselves with screening participation rates.

As before, I simulated screening for a synthetic cohort of one million men and one million women at age 50 at the start of the screening programme with follow-up over the rest of each population member's life. All participants were assumed to take part in all offers of screening. The results are presented in table 5.14, along with the corresponding results from FOBT screening without sigmoidoscopy. The addition of 5 yearly sigmoidoscopy to an annual FOBT screening programme results in the prevention of an extra 14% deaths from colorectal cancer, leading to an extra 15% fall in YLL.

5.8 Sensitivity analyses

Sensitivity analysis is a technique applied in simulation modelling to quantify the robustness of the models results to uncertainty in key modelling assumptions and parameters. In principle all the model's parameters and assumptions could be subject to sensitivity analysis, but the large number of such assumptions and parameters make this impractical. Instead I have focussed the sensitivity analyses on those model parameters which have the greatest uncertainty or which could be subject to variation as a policy decision in the design of a screening programme.

Table 5.15 lists the parameters subjected to sensitivity analysis and the alternative values tried for each parameter. Table 5.16 lists the cost per YLS and the fall in YLL for the standard model and each of the alternative parameters.

5.8.1 Disease model parameters

The disease model parameters included in the sensitivity analyses were the sojourn times in each cancer stage, the cure rates for stages A, B and C and the proportion of new polyps which will progress to cancer. The sojourn times were drawn from expert opinion with no empirical base. I tried simulating screening with each of the sojourn times arbitrarily increased by 50% and decreased by 50%. The distribution based model fitting for cure rates reported in chapter 3 produced lower estimated rates than the distribution free estimation procedure. I ran the simulation with the lower rate estimates. The estimate of the proportion of new polyps which would progress to cancer involved the assumption that none of the new polyps under 1 cm would progress. I tried screening with this parameter arbitrarily increased by 50%. The simulations with these alternative values all gave rise to changes in the cost per YLS of less than $1,000, with most less than $500. The changes is the simulated

mortality falls were mostly less than 1%—the exceptions being the variations in the stage A sojourn times and the cure rates which each led to changes in the mortality falls of around 1% to 1.5%. These are all differences which would be unlikely to lead to different screening programme policy decisions. Hence the model appears relatively robust to at least moderate changes in these parameters.

5.8.2 Screening model parameters

The screening model parameters included in the sensitivity analyses were the screening sensitivity, specificity and participation rates, both in the screening programme and in background screening. Screening sensitivity estimates play a key role in the model's results but, as noted in chapter 4, the published sensitivities from RCTs are not detailed enough for my model and in any case may not be reliable. The estimates from St. John & Young et al. (1993) are based on symptomatic patients and so are likely to be too high. I derived estimates from this study to use as an approximate upper bound on the true FOBT sensitivity. I used St. John & Young et al.'s sensitivity estimates for polyps directly. However, they did not publish separate estimates for cancers by stage. My model sensitivity estimates are roughly the same for large polyps and stage A cancer, so I took St. John & Young et al.'s large polyp estimate as the sensitivity for stage A cancer and their general cancer estimate as applying to stages B and C. As before, I assigned stage D cancers a sensitivity of 1.

Wagner et al. (1996) suggest values of 90% and 98% as a plausible range for FOBT test specificity, so this was the range I used in my sensitivity analyses. It is unlikely that any population based mass screening would achieve the high participation rates of the Minnesota Study, since it was based on a self-selected sample. However, the Funen study achieved higher participation rates than either the Nottingham or Burgundy studies apparently by having a more intensive recruitment strategy. I simulated the screening programme using the Funen participation rates to give an idea of the likely benefits of devoting more resources to recruitment.

There are no Australian data on background FOBT screening rates so I followed Salkeld et al. in using the rate found in the Minnesota study. I arbitrarily doubled this rate in the sensitivity analysis.

The higher sensitivity rates give rise to a greatly improved cost per YLS and a larger fall in mortality. The variation in specificity does not affect the estimated mortality fall but does substantially vary the cost of screening. This result shows the key role in the model played by the sensitivity and specificity estimates and highlights the importance of finding reliable estimates.

The higher screening programme participation rate does not substantially alter the cost per YLS but does increase the mortality fall by around 3.6%. The cost result should be treated with caution because my costs do not include recruitment costs which could be much higher with more intensive recruitment strategies. However, the extra fall in mortality with a recruitment rate that has been achieved in population screening overseas and so is potentially achievable here highlights the

need to carefully consider recruitment strategies in planning a screening programme.

The higher rate of background screening has no effect on mortality because the model uses the observed mortality for the mortality in the absence of screening and only uses the background screening in cost calculations. Doubling the rate changes the cost per YLS by less than $300, suggesting the model is fairly robust to errors in this parameter.

5.8.3 Screening costs and benefits

The parameters effecting screening costs and benefits which were included in the sensitivity analyses were the discount rate, the cost of FOBT screening, the cost of colonoscopy follow up and the cost of cancer treatment.

I ran the simulation with a 5% annual discount rate on both costs and YLS as suggested by the US Panel for Cost-Effectiveness in Health and Medicine (Weinstein et al. 1996) and with a 10% annual discount rate as suggested by Salkeld et al. Salkeld et al. also suggested a plausible range for the cost of FOBT screening would be $11 to $21 in 1994, which translates to a range of $9.60 to $18.40 in 1990. I also arbitrarily decreased and increased the cost of colonoscopy follow up by 50% and arbitrarily increased the cost of cancer treatment by 50%.

Increasing the cost of FOBT screening increased the cost per YLS by around $500 while decreasing the FOBT screening cost decreased the cost per YLS by around $1000. This suggests that the screening programme costs are relatively insensitive to changes in the cost of FOBTs. On the other hand, increasing the costs of colonoscopy follow up increased the cost per YLS by around $5,000 while decreasing the cost of colonoscopy decreased the cost per YLS by a similar amount. This suggests that controlling the cost of colonoscopy follow up would have a bigger impact on the cost of a FOBT screening programme than controlling the cost of the FOBT screen.

Increasing the cost of cancer treatment by 50% reduced the cost per YLS by around $2,500. This is because screening both prevents some cancers and leads to earlier stage detection, and lower cost treatment, in others. Hence a greater treatment cost leads to a bigger difference in costs between a population with and without screening.

Salkeld et al. found similar results when applying sensitivity analyses to the discount rate and screening and treatment costs.

5.9 Tables

Table 5.1: Results from ten simulations of the base model applied to the whole Australian population

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
Maximum value for the ten simulations	$7,836	$40,096	20.6%	22.4%
Minimum value for the ten simulations	$7,669	$39,378	20.3%	22.1%
Difference (maximum - minimum)	$167	$719	0.3%	0.3%

Table 5.2: Results for biennial screening base model and alternative screening protocols

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
Base model—biennial screening, ages 45-74	$7,690	$39,499	20.5%	22.2%
Biennial screening, ages 40-74	$8,895	$46,601	22.1%	23.5%
Biennial screening, ages 50-74	$6,904	$34,610	19.5%	21.7%
Biennial screening, ages 55-74	$5,898	$27,850	16.2%	19.2%
Biennial screening, ages 45-69	$8,146	$45,249	17.5%	17.6%
Biennial screening, ages 45-79	$7,586	$36,768	22.3%	25.6%
Biennial screening, ages 45-84	$7,613	$35,515	23.1%	27.6%
Biennial screening, ages 45 and over	$7,853	$35,873	23.5%	28.6%
Annual screening, ages 45-74	$9,779	$50,850	26.6%	28.6%

Table 5.3: Results for annual screening base model and alternative screening protocols

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
Base model—annual screening, ages 50-74	$8,793	$44,397	25.4%	28.1%
Annual screening, ages 45-74	$9,779	$50,850	26.6%	28.6%
Annual screening, ages 40-74	$11,317	$60,033	28.3%	29.8%
Annual screening, ages 55-74	$7,535	$36,051	21.5%	25.0%
Annual screening, ages 50-69	$9,038	$48,895	21.8%	22.5%
Annual screening, ages 50-79	$8,828	$42,243	27.5%	32.1%
Annual screening, ages 50-84	$8,987	$41,777	28.5%	34.2%
Annual screening, ages 50 and over	$9,328	$42,417	28.7%	35.2%

Table 5.4: Results from ten simulations of annual screening between the ages of 50 and 84 for the synthetic cohort aged 50 at the start of screening

	% fall in deaths due to screening	% fall in YLL due to screening
Mean	67.9%	62.0%
Highest value	68.2%	62.3%
Lowest value	67.5%	61.5%
Difference (highest - lowest)	0.8%	0.8%

Table 5.5: Simulated falls in mortality due to annual screening up to age 84, for people taking part in the screening program starting at selected ages

Age at start of screening	% fall in deaths due to screening	% fall in YLL due to screening
50	68.0%	62.1%
70	52.2%	45.0%

Table 5.6: Average increase in years of life lived for each true positive screen, for annual screening between up to age 84, starting at selected ages

Age at start of screening	Average increase in years of life lived
50	2.1
70	1.7

Table 5.7: Cost per YLS of selected health interventions at 1989-90[a] prices

Program	Cost per life year[b]
Health promotion program	
Sydney Quit Smoking Campaign	$16
Non-drug blood pressure reduction clinic	$5,163
Care/cure programs	
Neonatal intensive care, babies 1,000–1,500 kg	$1,240–3,099
Kidney transplant	$4,719
Neonatal intensive care, babies < 801 kg	$3,718–4,751
Breast cancer screening[c]	$6,816–11,360
Colorectal cancer screening	$10,912
Cervical cancer screening[c]	$31,790
Hospital dialysis	$49,354
AIDS treatment with zidovudine	$134,257

Notes:

(a) Costs adjusted to 1989–90 prices using the AIHW total health price index (AIHW 1997).

(b) All cost per life year estimates calculated with 5% per year discount rate.

(c) Cost per life year estimates for breast and cervical cancer screening programs are the estimates calculated before the programs were implemented—estimates based on subsequent experience with these programs may produce different values.

Source: AHMAC (1991)

Table 5.8: Average cost per YLS and per death averted for annual FOBT screening at ages 50 to 84 with marginal costs associated with changing the starting and finishing screening ages and moving to biennial screening

	Cost per life year saved	Cost per death averted
Annual FOBT screening at ages 50 to 84	$8,987	$41,777
	Cost per extra life year saved	**Cost per extra death averted**
Moving start age from 50 to 45	$31,751	$494,885
Moving start age from 55 to 50	$15,526	$111,303
Moving end age from 84 to death	$51,111	$63,305
Moving end age from 79 to 84	$13,630	$34,539
Moving from biennial to annual screening	$16,721	$90,687

Table 5.9: Sensitivity and specificity for rehydrated FOBT

Test parameter	(%)
Sensitivity	
Polyp less than 5 mm	2.5
Polyp 5 to 9 mm	3.4
Polyp 10 to 19 mm	12.6
Polyp 20 mm or more	22.8
Cancer stage A	22.8
Cancer stage B	51.4
Cancer stage C	67.3
Cancer stage D	100.0
Specificity	87.9

Source: Derived from Mandel et al. (1993).

Table 5.10: Results for annual screening between the ages of 50 and 84 for rehydrated and unrehydrated FOBT

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Unrehydrated slides	$8,987	$41,777	28.5%	34.2%
Rehydrated slides	$16,538	$77,398	30.5%	36.4%
Marginal cost per extra death averted or YLS	$118,900	$612,027		
Difference in mortality fall			2.0%	2.2%

Table 5.11: Sensitivity, specificity and cost for immunochemical FOBT

Test parameter	
Sensitivity	(%)
Polyp less than 5 mm	4.8
Polyp 5 to 9 mm	6.5
Polyp 10 to 19 mm	19.7
Polyp 20 mm or more	35.8
Cancer stage A	48.1
Cancer stage B	84.3
Cancer stage C	84.3
Cancer stage D	100.0
Specificity	97.8
Cost ($A1990)	$18.40

Source: Polyp sensitivity derived from St. John & Young et al. (1993)
Cancer sensitivity directly from Shimboet al. (1994)
Specificity directly from St. John & Young et al. (1993)
Cost derived from Salkeld et al. (1996).

Table 5.12: Results for annual screening between the ages of 50 and 84 for standard guaiac and immunochemical FOBT

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Standard guaiac FOBT	$8,987	$41,777	28.5%	34.2%
immunochemical FOBT	$7,629	$36,204	38.7%	45.5%
Difference in mortality fall			10.2%	11.3%

Table 5.13: Results for annual FOBT screening between the ages of 50 and 84 with sigmoidoscopy offered every five years

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Standard FOBT screening without sigmoidoscopy	$8,987	$41,777	28.5%	34.2%
Screening where all participants accept sigmoidoscopy	$6,822	$32,431	39.4%	46.3%
Screening with sigmoidoscopy participation rate of 15%	$8,158	$38,186	31.1%	37.1%

Table 5.14: Simulated falls in mortality due to annual FOBT screening up to age 84, for people taking part in the screening program starting at age 50

	% fall in deaths due to screening	% fall in YLL due to screening
Screening without sigmoidoscopy	68.0%	62.1%
Screening with sigmoidoscopy	81.8%	77.5%

Table 5.15: Model parameters subject to sensitivity analysis

Model parameter	Model value	Lower Alternate value	Upper Alternate value
Disease model			
Sojourn time increased and decreased by 50%			
stage A (years)	2.00	1.00	3.00
stage B(years)	1.00	0.50	1.50
stage C(years)	1.00	0.50	1.50
stage D(years)	0.75	0.375	1.125
Lower cure rates			
stage A (%)	78.5	64.0	—
stage B (%)	57.1	46.7	—
stage C (%)	30.0	25.9	—
Proportion of new polyps proceeding to cancer increased by 50% (%)	14	—	21

Table 5.15: Model parameters subject to sensitivity analysis *(continued)*

Model parameter	Model value	Lower Alternate value	Upper Alternate value
Screening model—FOBT screening			
Higher sensitivity, derived from St.John & Young et al. (1993) (%)			
polyps <5 mm	2	—	16.7
polyps 5-9 mm	3	—	16.7
polyps 10-19 mm	11	—	42.2
polyps 20+ mm	20	—	42.2
Cancers stage A	20	—	42.2
Cancers stage B	45	—	88.8
Cancers stage C	59	—	88.8
Cancers stage D	100	—	100
Specificity from Wagner et al. (1996) (%)	95	90	98
Higher participation rates, from Funen study (Kronborg et al. 1996) (%)			
Males			
Ages under 50 years	34	—	76
Ages 50-54 years	51	—	73
Ages 55-59 years	52	—	72
Ages 60-64 years	54	—	71
Ages 65-69 years	54	—	64
Ages 70 years and over	49	—	50
Females			
Ages under 50 years	43	—	76
Ages 50-54 years	59	—	73
Ages 55-59 years	59	—	72
Ages 60-64 years	58	—	71
Ages 65-69 years	54	—	64
Ages 70 years and over	47	—	50
Background screening participation rate doubled (%)	1.8	—	3.6
Costs and benefits—FOBT screening			
Costs and YLL annual discount rate (%)	3	—	5
Costs and YLL annual discount rate (%)	3	—	10
Cost of FOBT screening derived from Salkeld et al.(1996) ($A 1990)	14.00	9.60	18.40
Cost of colonoscopy follow up increased & decreased by 50% ($A 1990)	715.00	357.50	1072.50
Cost of cancer treatment (increased by 50% $A 1990)			
Stage A	13,324	—	19,986
Stage B	14,462	—	21,693
Stage C	14,631	—	21,947
Stage D	18,319	—	27,479
Advanced	21,300	—	31,950

Table 5.16: Results of sensitivity analyses

Model parameter	Cost per YLS	Per cent fall in YLL due to screening
Annual FOBT, ages 50 to 84 from standard model	$8,987	28.5%
Disease model		
Lower sojourn time		
stage A	$9,660	29.6%
stage B	$9,417	29.0%
stage C	$9,414	28.4%
stage D	$9,260	28.1%
Higher sojourn time		
stage A	$8,462	27.1%
stage B	$8,734	27.6%
stage C	$8,576	28.5%
stage D	$8,447	28.9%
Lower cure rates	$8,468	27.0%
Higher proportion of new polyps proceeding to cancer	$8,588	28.4%
Screening model—FOBT screening		
Higher sensitivity	$5,380	43.7%
Higher specificity f	$5,275	N/A
Lower specificity	$15,029	N/A
Higher participation rates	$8,888	32.1%
Higher background screening participation	$8,726	N/A
Costs and benefits—FOBT screening		
Costs and YLL 5% annual discount rate	$10,588	N/A
Costs and YLL 10% annual discount rate	$18,473	N/A
Lower cost of FOBT screening	$7,937	N/A
Higher cost of FOBT screening	$9,515	N/A
Lower cost of colonoscopy follow up	$3,604	N/A
Higher cost of colonoscopy follow up	$13,848	N/A
Higher cost of cancer treatment	$6,423	N/A

5.10 Figures

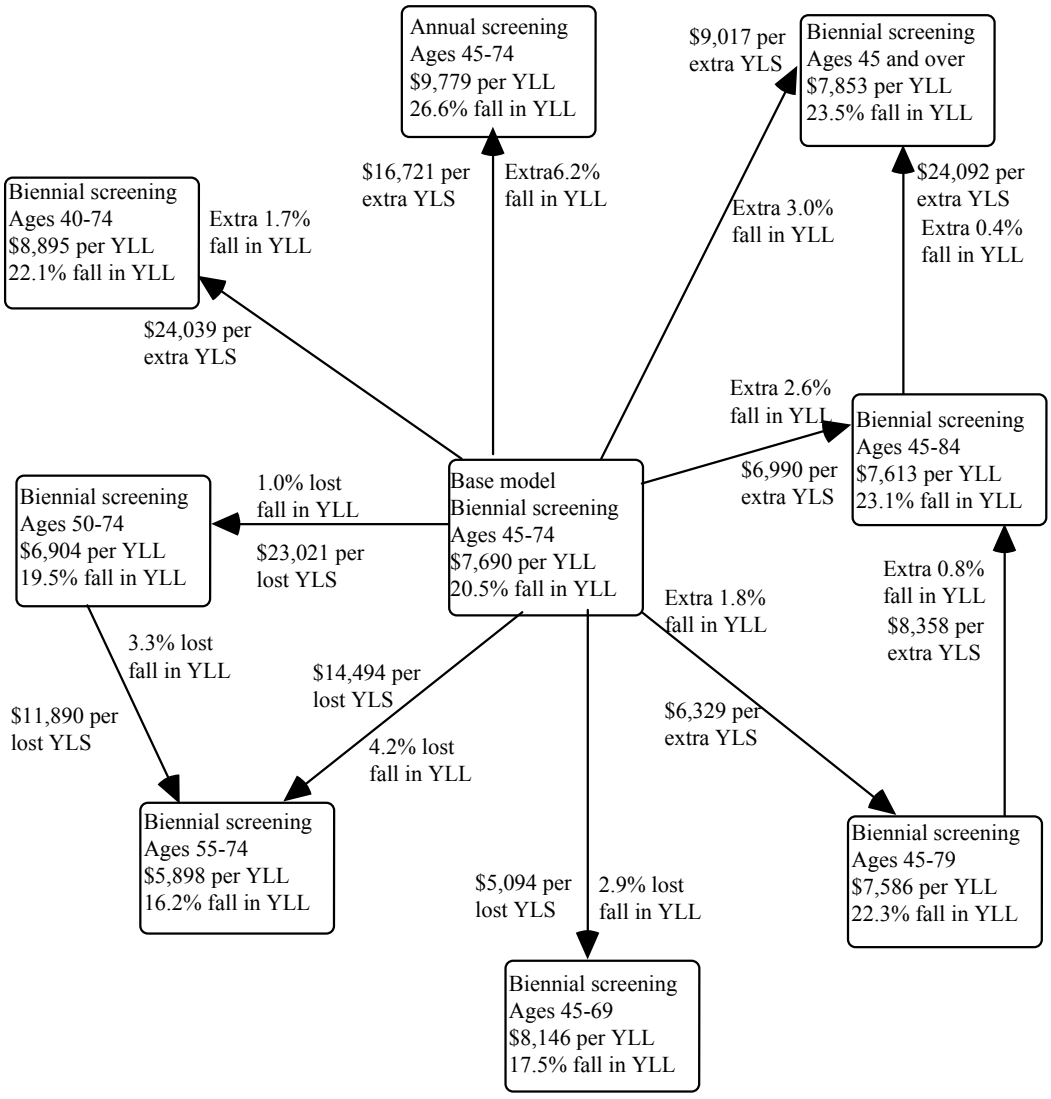

Figure 5.1: Change in YLL and cost per extra life year saved when moving between biennial screening base model and alternative screening protocols

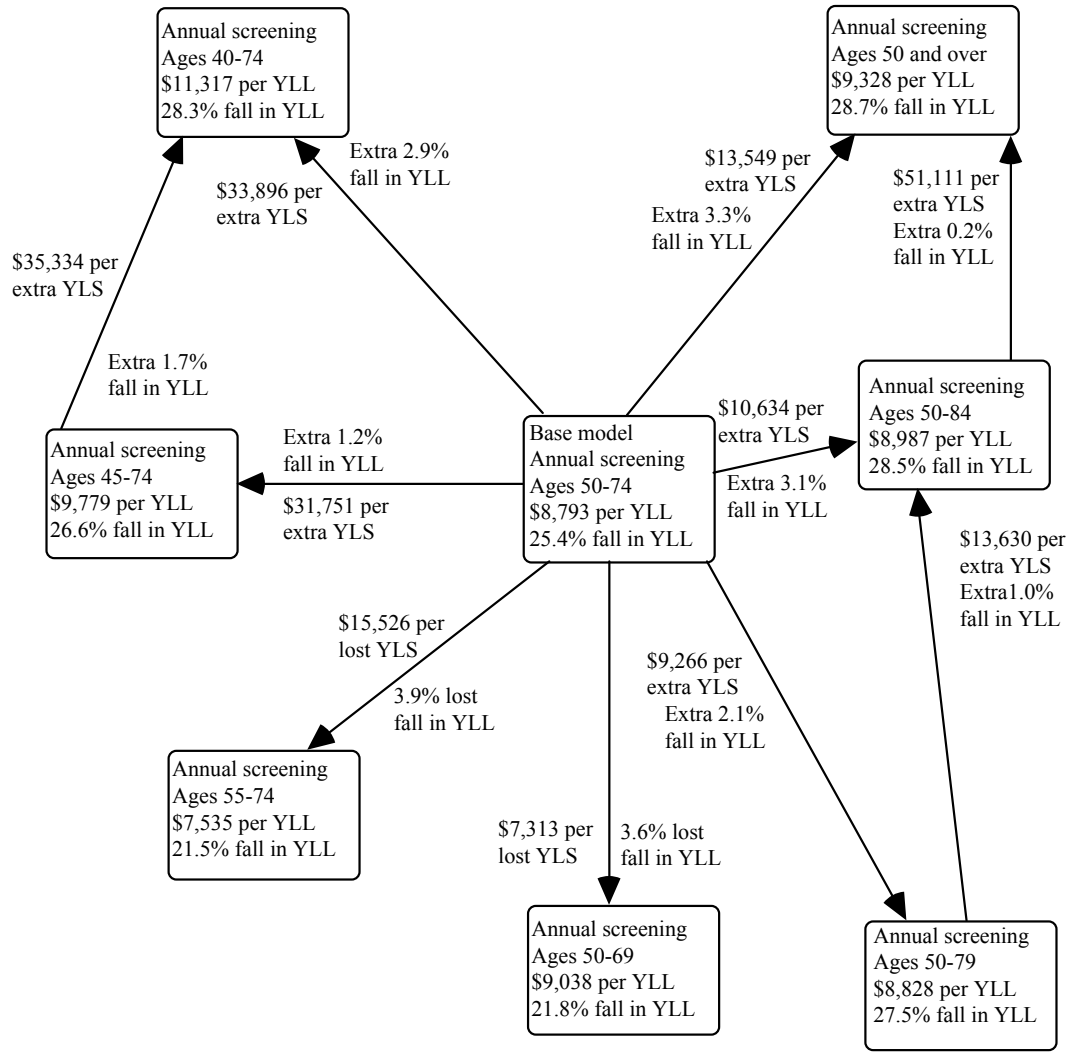

Figure 5.2: Change in YLL and cost per extra life year saved when moving between annual screening base model and alternative annual screening protocols

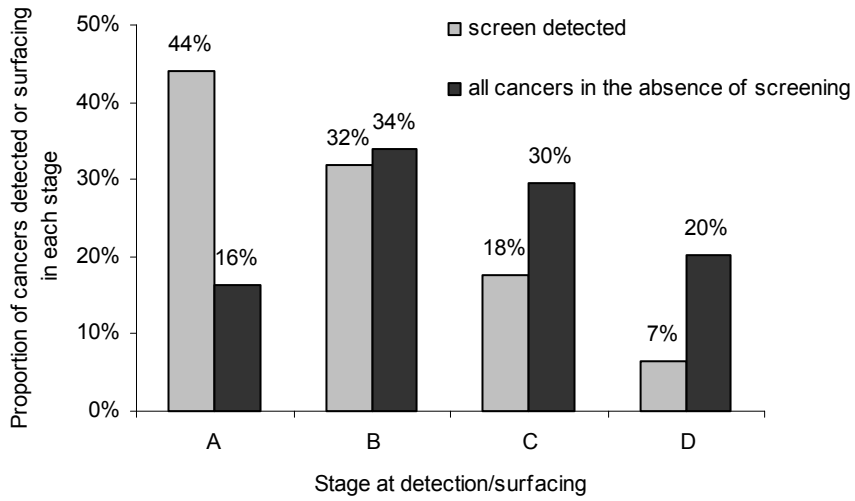

Figure 5.3: Simulated distribution of cancers by stage at detection or clinical surfacing for annual screening of the Australian population between the ages of 50 and 84

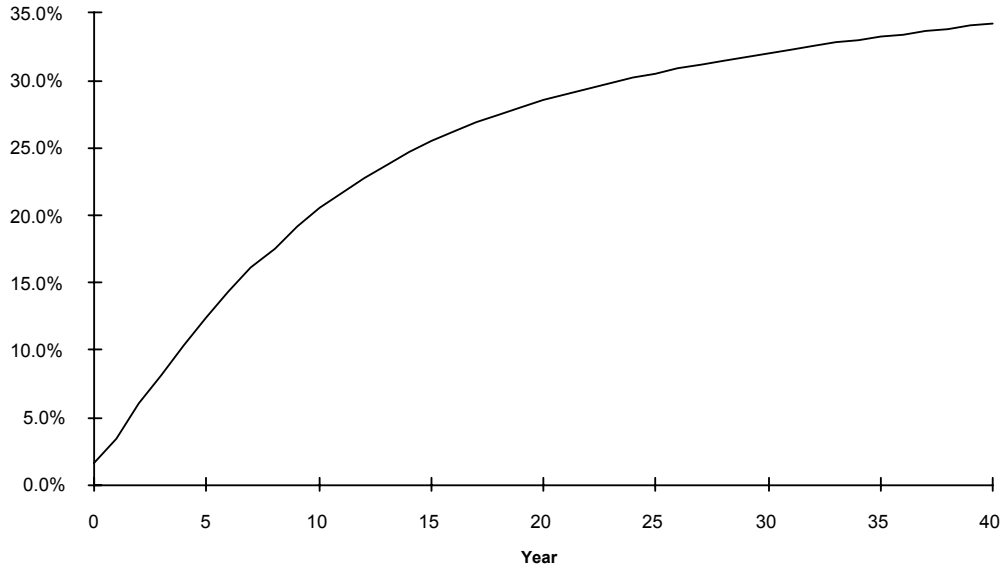

Figure 5.4: Simulated cumulative colorectal cancer mortality fall by year for annual screening of the Australian population between ages 50 and 84

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

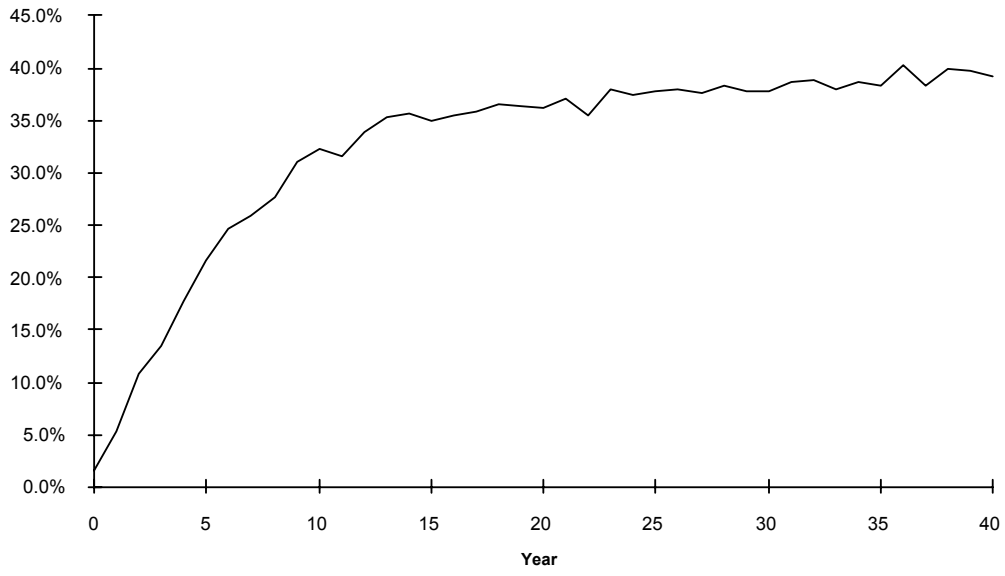

Figure 5.5: Simulated annual colorectal cancer mortality fall by year for annual screening of the Australian population between ages 50 and 84

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

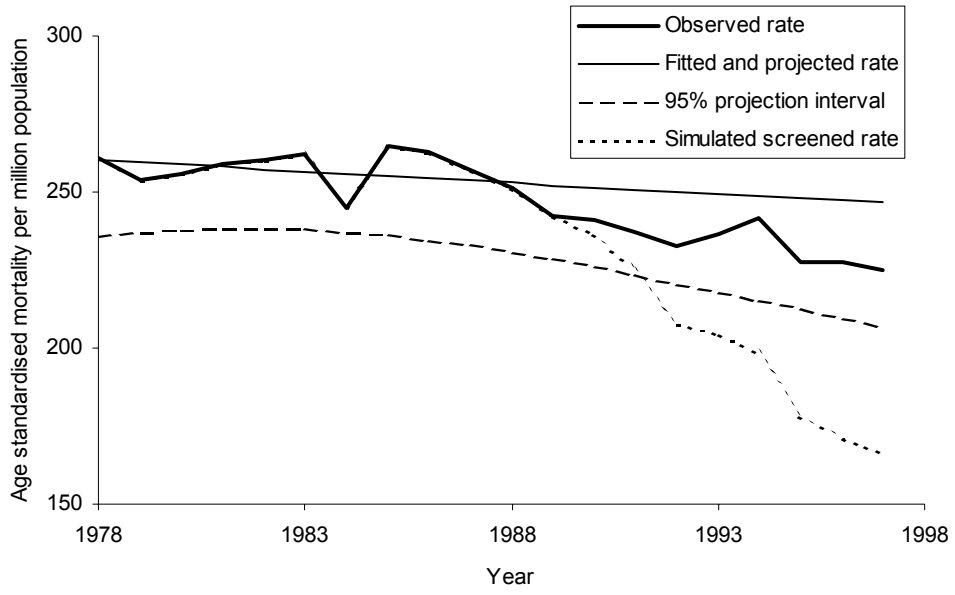

Figure 5.6: Comparison of projected colorectal cancer mortality without screening and simulated mortality with screening

Notes: 1. All mortality rates are age standardised to the total 1988 Australian population and calculated per million population.
2. The fitted and projected rates are calculated from a poisson regression model fitted to mortality rates from 1972 to 1989 inclusive.
3. The screened rates are calculated by applying the simulated mortality fall due to screening to the actual mortality rates for the period 1990 to 1997.

6 Screening with other modalities

6.1 Introduction

The major studies of population screening for colorectal cancer have focussed on FOBT testing with some other follow up of positive results. However, while both sigmoidoscopy and colonoscopy have been suggested as primary screening tools (Winawer et al. 1997), there have been no major studies of population screening programmes using these modalities. This chapter presents the results of simulations of a population screening programme using each of these as the primary screening tool.

Because neither colonoscopy nor sigmoidoscopy have been widely studied as a tool for mass population screening, the likely response of the general population is less well understood than for FOBT screening. I have derived a screening response rate from studies of population screening with sigmoidoscopy, but this may be too pessimistic in practice. Hence I will examine more optimistic but still plausible screening response rates as part of my sensitivity analysis.

As with FOBT screening, cancers identified by screening are treated identically to cancers arising in the absence of screening. Polyps larger than 1 cm identified by screening are removed and the person followed up by colonoscopy at three yearly intervals. Polyps smaller than 1 cm are removed and the person goes back into the screening programme. The parameters governing the simulation will be those listed in table 4.1.

The analyses presented here start with a base model and investigate the effect of varying the screening interval. Once I have established a preferred screening interval, I will vary the ages at which screening is offered to find a preferred target age range. Then I will examine the benefits of screening for individuals who take part in the screening programme using my preferred screening interval and target age range. Finally I will examine the increase in benefits over time from the introduction of this screening programme.

The analysis of screening benefits to programme participants will be done by simulating screening for synthetic cohorts starting at the youngest age for my screening protocol and starting at age 70, with follow-up over the whole of the rest of each population member's life. The aim is to examine benefits both to those who participate in the whole screening programme and those who by commencing screening at a later age only participate in part of it.

I will examine screening benefits over time by first plotting the simulated fall in both annual and cumulative mortality in the 40 years following the beginning of the population screening programme. I will then project colorectal cancer mortality

from 1990 to 1997 using a poisson regression model as described in chapter 5. I will then apply the simulated annual mortality falls from my model to the actual mortality rates for the period 1990 to 1997 and compare these with a one sided 95% projection interval from the poisson regression. This should show when mortality falls from a screening programme will be detectable.

There are no studies of population screening programmes using colonoscopy or sigmoidoscopy against which this screening simulation model can be validated—the Australian study which was the basis of the screening participation rates did not present its results in a way which allowed direct simulation. Hence the results presented here should be treated with caution. However, this model is based on the best available information on screening with colonoscopy or sigmoidoscopy, so the simulation results represent the best available indication of its costs and benefits.

6.2 Screening with colonoscopy

This section will examine screening using colonoscopy. That is, I will assume that each participant is offered a colonoscopy at regular intervals. Where colonoscopy is incomplete, the addition of a double contrast barium enema is used to screen the entire colon.

One difficulty with colonoscopy as a tool for mass screening is the need for suitably trained and experienced physicians for its administration. Further problems would be likely to arise in ensuring not only adequate numbers of trained practitioners but also their geographic distribution so that all Australians would have access to such screening. Consideration of the difficulty and expense involved in recruiting, training and equipping sufficient suitably qualified practitioners to administer a population screening programme is beyond the scope of this study. However, the experience of the Australian mammography screening programme suggests that mobilising sufficient trained medical practitioners to administer medically sophisticated screening techniques is possible but would need careful planning and adequate resourcing.

In this chapter I will take as my initial base case the screening protocol developed for FOBT screening. That is each person aged between 50 and 84 will be offered screening. Other modelling studies suggest that annual screening is too short a screening interval for colonoscopy (Wagner et al. 1996), so I will take three years as the interval for my initial base model and compare this with screening the same age group but with longer intervals.

6.2.1 Population based mass screening

Screening interval

Table 6.1 lists the percentage fall in cumulative mortality and cumulative discounted life years lost and the cost per life year saved and per death averted for ten simulations of the base model applied to the whole Australian population. There

is less than half a per cent difference between the highest and lowest values of the in cumulative deaths prevented and YLS. The highest and lowest cost per life year saved differ by less than $300 while the highest and lowest cost per death prevented differ by around $1,500. While this shows higher variability than the model for FOBT explored in chapter 4, the differences between screening protocols in each of these measures would still have to be greater than this to have any practical implications.

Figure 6.1 shows the average cost per year of life saved at each screening interval from one to 15 years. These show a steep fall in average cost per YLS for each year added on to the screening interval up to around 7 years. After 10 years, each year added to the screening interval changes the average cost per YLS by less than $400.

Table 6.2 shows the average cost per YLS and death averted and the average percentage fall in YLL and deaths due to screening at each interval from 3 to 11 years. Figure 6.2 shows the extra YLS and cost per YLS in moving between the different screening intervals. For these screening intervals there is less than 1% increase in YLS when the interval is reduced by a year. However, the extra cost per YLS drops substantially from nearly $750,000 for a screening interval decrease from 4 years to 3 years to around $26,000 for a decrease from 11 years to 10 years. Changing the interval from 10 years to 9 years costs around $39,500 for each extra YLS, which is above the limit of $32,000 suggested by George et al (1998), while changing it from 11 years to 10 years is below this limit. This suggests that a screening interval of 10 years is appropriate for mass screening using colonoscopy as the primary screening tool.

Screening age range

Table 6.3 shows the average cost per YLS and death averted and the average percentage fall in YLL and deaths due to screening at different age ranges. Figure 6.3 shows the extra YLS and cost per YLS in moving between the different age ranges. Moving from 50 to 45 as the starting age gives a substantial gain in YLS at a cost per YLS of around $27,500, which is well below the $32,000 limit. Moving from 40 to 35 gives around half the gain in YLS at a cost of around $58,500 per YLS, which is well above the limit. However, the choice between 45 and 40 is not so clear. Moving from 45 to 40 as the starting age gives a moderate gain in YLS at a cost of around $35,000 per YLS. This is above my limit of $32,000, but only by $3,000. Further, the cost per extra death averted in moving from 45 to 40 is around $243,000, which is below the limit of $260,000. The $32,000 per YLS limit is only a rough guide, so the move from 45 to 40 can be seen as on the borderline of acceptability in cost-effectiveness terms. I chose to limit screening to ages 45 and over for the purposes of this analysis, but a policy maker could justifiably choose to start at age 40 and would probably make the choice between these two ages on grounds other than cost-effectiveness.

Moving from 79 to 84 as the finishing age leads to a small increase in YLS at a cost of around $30,000 per extra YLS. Moving from stopping screening at age 84 to

having no upper age limit leads to a very small increase in YLS at a cost of around $93,000 per extra YLS.

These results suggest that colonoscopy screening every 10 years starting at age 45 and finishing at age 84 would be an acceptable mass screening programme in cost-effectiveness terms. Obviously in the long term people entering such a programme at age 45 would have their last screen at age 75 and, in the absence of the upper age limit, be eligible for their next screen at age 85. Hence, an upper age limit of 85 could be more appropriate than 84. My study is focussed on the first 40 years of a screening programme. The first cohort who enter the programme at age 45 would only reach the upper age limit at the end of my study period, so this issue would not affect my results. However, since a programme of screening at 10 year intervals between ages 45 and 85 is more likely to be implemented than one between ages 45 and 84, this is the programme I will choose as my preferred screening protocol. The average cost per YLS and per death averted and the proportion of YLL and deaths prevented for this age range are also shown in table 6.3. The proportion of screen detected cancers in each stage are presented in figure 6.4, along with the proportion of all cancers clinically detected in each stage in the absence of screening.

6.2.2 The benefits of colonoscopy screening for screening participants

Table 6.4 lists the percentage fall in cumulative mortality and YLL due to screening for ten simulations of my synthetic cohort aged 45 years at the start of screening. The simulated values vary by 0.5% for falls in mortality and by 0.4% for falls in YLL across the 10 simulations.

Table 6.5 lists the simulated falls in both deaths and discounted YLL due to screening for screening starting at each of the selected ages. For the cohort aged 45 at the start of screening the average risk of death from colorectal cancer over the rest of their life fell by 79%, while the average years of life lost due to colorectal cancer fell by 76%. For the cohort which entered the screening programme at age 70 the average risk of death from colorectal cancer over the rest of their life fell by 74%, while the average years of life lost due to colorectal cancer fell by 69%.

Table 6.6 lists the average years of life gained for each true positive screen. Each positive screen resulted in an average of 2.2 extra years of life for people starting screening at age 45 and an extra 2.0 years of life for people starting screening at age 70. These are slightly higher than the corresponding values for FOBT screening because colonoscopy detects cancers at an earlier stage than FOBT (including more in the polyp stage), resulting in a higher cure rate.

6.2.3 The benefits of colonoscopy screening over time

Figure 6.5 shows the fall due to screening in cumulative mortality from colorectal cancer by year from the introduction of screening and figure 6.6 shows the corresponding fall due to screening in annual mortality from colorectal cancer. Both the cumulative and annual mortality gains rise sharply for the first 10 to 15 years,

the cumulative to around 28% and the annual to around 45%. They both continue rising more slowly to the end of the 40 year screening period to reach values of 47% and 60% respectively.

The results of the mortality projection using poisson regression are summarised in figure 6.7. These suggest that the mortality rate with screening would become statistically significantly different from the projected rate during the third year of screening.

This means that, as with FOBT screening, a national monitoring system should be able to demonstrate that screening was having an effect on mortality after three years of a screening programme. However, the simulation suggests that while substantial gains can be demonstrated after 10 to 15 years of screening, the programme would not achieve its full potential gains before the end of the 40 year screening period.

6.3 Screening with sigmoidoscopy

This section will examine screening using sigmoidoscopy. That is, I will assume that each participant is offered a sigmoidoscopy at regular intervals and those with a positive result for any of the smears are offered a complete bowel examination with a colonoscopy.

Like colonoscopy, sigmoidoscopy needs suitably trained and experienced health professionals for its administration. However mobilising sufficient trained health professionals presents less difficulty with sigmoidoscopy than with colonoscopy. All screening in the Australian sigmoidoscopy screening study was carried out at Freemantle hospital (Olynyk et al. 1996), but expert opinion suggests that GP's can be trained to administer sigmoidoscopy screening (Graham 1996) (this article focuses on rigid sigmoidoscopy but notes that GP's can be trained to use the longer flexible sigmoidoscopes expertly). Further, overseas studies show that non physician health professionals can be trained to administer sigmoidoscopic screening (Winawer et al. 1997).

In this section I will take as my initial base case the screening protocol developed for colonoscopy screening—that is offering screening to each person aged between 45 and 85. I will take annual screening as the interval for my initial base model.

6.3.1 Population based mass screening

Screening interval

Table 6.7 lists the percentage fall in cumulative mortality and cumulative discounted life years lost and the cost per life year saved and per death averted for ten simulations of the base model applied to the whole Australian population. There is half a per cent or less difference between the highest and lowest values in cumulative deaths prevented and YLS. The highest and lowest cost per life year

saved differ by less than $150 while the highest and lowest cost per death prevented differ by around $500. The differences between screening protocols in each of these measures would have to be greater than this to have any practical implications.

Table 6.8 shows the average cost per YLS and death averted and the average percentage fall in YLL and deaths due to screening at each interval from 1 to 6 years. Each of these screening intervals is associated with an average cost per YLS which is well below my $32,000 limit. Figure 6.8 shows the extra YLS and cost per YLS in moving between the different screening intervals. For screening intervals up to three years there is less than 1% increase in YLS when the interval is reduced by a year. However, the extra cost per YLS in moving from two years to one year is almost $1.5 million and in moving from three to two years is around $63,000, both of which are well above my limit. When I move from four years to three years I get a bigger reduction in YLL with an extra cost per YLS of around $27,000, which is well below my limit. This suggests that a screening interval of 3 years is appropriate for mass screening using sigmoidoscopy as the primary screening tool.

Screening age range

Table 6.9 shows the average cost per YLS and death averted and the average percentage fall in YLL and deaths due to screening at different age ranges. Again, all of the age ranges are associated with an average cost per YLS which is well below my $32,000 limit. Figure 6.9 shows the extra YLS and cost per YLS in moving between the different age ranges. Moving from 45 to 40 as the starting age gives a substantial gain in YLS at a cost per YLS of around $11,300, which is well below the $32,000 limit. Moving from 40 to 35 gives a little more than half the gain in YLS at more than double the cost (around $23,300 per YLS). On strict cost-effectiveness grounds this is still an acceptable cost. However, cancers occurring below age 40 are mainly associated with one of the high risk groups which are likely to be already under treatment (Winawer et al. 1997) and hence outside the scope of a mass screening programme. So I chose to limit screening to people aged 40 and over.

Moving from 80 to 85 as the finishing age leads to a small increase in YLS at a cost of around $23,000 per extra YLS. Further, this small increase in YLS is associated with a moderate increase in deaths averted (1%). Moving from stopping screening at age 85 to having no upper age limit leads to a very small increase in YLS at a cost of around $45,000 per extra YLS. These results suggest that sigmoidoscopy screening every 3 years starting at age 40 and finishing at age 85 would be an acceptable mass screening programme in cost-effectiveness terms. The proportion of screen detected cancers in each stage for this programme are presented in figure 6.10, along with the proportion of all cancers clinically detected in each stage in the absence of screening.

6.3.2 The benefits of sigmoidoscopy screening for screening participants

Table 6.10 lists the percentage fall in cumulative mortality and YLL due to screening for ten simulations of my synthetic cohort aged 40 years at the start of screening. The simulated values vary by 0.5% for falls in mortality and by 1.1% for falls in YLL across the 10 simulations.

Table 6.11 lists the simulated falls in both deaths and discounted YLL due to screening for screening starting at each of the selected ages. For the cohort aged 40 at the start of screening both the average risk of death from colorectal cancer over the rest of their life and the average years of life lost due to colorectal cancer fell by around 69%. For the cohort which entered the screening programme at age 70 the average risk of death from colorectal cancer over the rest of their life fell by 50%, while the average years of life lost due to colorectal cancer fell by 45%.

Table 6.12 lists the average years of life gained for each true positive screen. Each positive screen resulted in an average of 2.1 extra years of life for people starting screening at age 40 and an extra 1.9 years of life for people starting screening at age 70.

6.3.3 The benefits of sigmoidoscopy screening over time

Figure 6.11 shows the fall due to screening in cumulative mortality from colorectal cancer by year from the introduction of screening and figure 6.12 shows the corresponding fall due to screening in annual mortality from colorectal cancer. The cumulative mortality gains rise fairly steadily across the whole 40 year period, reaching the value of 38% by the end of the period. The annual mortality gains rise fairly steadily over the first 18 to 20 years to a value of around 40%. After this the rise is slower, reaching a plateau of around 50% after 35 years

The results of the mortality projections from the poisson regression are summarised in figure 6.13. These suggest that, unlike FOBT and colonoscopy screening, the mortality rate with screening would become statistically significantly different from the projected rate during the sixth year of screening.

This means that a national monitoring system should be able to demonstrate that screening was having an effect on mortality after six years of a screening programme. However, the simulation suggests that while substantial gains can be demonstrated after around 20 years of screening with the programme achieving its full potential gains after around 35 years.

6.4 Sensitivity analyses

Two model parameters were subject to sensitivity analyses—the cost of either sigmoidoscopy or colonoscopy as a screening tool and the screening participation rates. The preferred screening programme was simulated with the cost of

colonoscopy or sigmoidoscopy increased and decreased by 50% and with the participation rates set to the same as for FOBT screening. The results are presented in table 6.13, along with the results from the simulation using the standard model for comparison.

Changing the cost of the screen by 50% substantially changes the cost per YLS, by around $10,000 for colonoscopy and by around $3,000 for sigmoidsocopy. This suggests, not surprisingly, that unlike FOBT screening, substantial gains could be made in overall screening programme costs by controlling the cost of the primary screening tool.

Increasing the participation rate to that of FOBT screening substantially increases the mortality gains by around 10.5% for colonoscopy and 8.5% for sigmoidoscopy. The simulated cost per YLS doesn't change appreciably, but the model does not include recruitment costs so this does not really represent the true costs of increasing screening participation. However, this results does suggest that putting resources into recruitment to increase screening participation rates could be a worthwhile investment.

6.5 Tables

Table 6.1: Results from ten simulations of the colonoscopy base model applied to the whole Australian population

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
Maximum value for the ten simulations	$27,715	$119,900	37.6%	48.7%
Minimum value for the ten simulations	$27,483	$118,328	37.3%	48.3%
Difference (maximum - minimum)	$231	$1,572	0.4%	0.4%

Table 6.2: Results for colonoscopy screening from ages 50 to 84 at different screening intervals

Screening interval	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
3 years	$27,643	$118,893	37.5%	48.5%
4 years	$22,455	$96,873	37.2%	48.2%
5 years	$19,307	$83,130	36.6%	47.4%
6 years	$17,431	$74,738	35.8%	46.6%
7 years	$16,106	$69,003	35.1%	45.7%
8 years	$15,115	$64,645	34.4%	44.9%
9 years	$14,401	$61,622	33.7%	43.8%
10 years	$13,852	$59,214	33.0%	42.9%
11 years	$13,534	$57,751	32.2%	42.0%

Table 6.3: Results for colonoscopy screening at ten year intervals for different age ranges

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
ages 50-84	$13,852	$59,214	33.0%	42.9%
ages 45-84	$15,472	$69,019	37.4%	46.8%
ages 40-84	$17,189	$79,184	40.9%	49.5%
ages 35-84	$19,280	$90,288	43.0%	51.3%
ages 55-84	$12,872	$51,918	26.9%	37.2%
ages 50-79	$13,341	$58,222	32.0%	40.7%
ages 50 and over	$14,546	$61,284	33.2%	43.9%
ages 45-85	$15,585	$69,246	37.6%	47.1%

Table 6.4: Results from ten simulations of colonoscopy screening every ten years between the ages of 45 and 85 for the synthetic cohort aged 45 at the start of screening

	% fall in deaths due to screening	% fall in YLL due to screening
Mean	79.3%	76.1%
Highest value	79.6%	76.4%
Lowest value	79.0%	75.9%
Difference (highest - lowest)	0.5%	0.4%

Table 6.5: Simulated falls in mortality due to colonoscopy screening every ten years up to age 85, for people taking part in the screening program starting at selected ages

Age at start of screening	% fall in deaths due to screening	% fall in YLL due to screening
45	79.4%	76.1%
70	73.5%	66.8%

Table 6.6: Average increase in years of life lived for each true positive screen, for colonoscopy screening every ten years up to age 85 starting at selected ages

Age at start of screening	Average increase in years of life lived
45	2.2
70	2.0

Table 6.7: Results from ten simulations of the sigmoidoscopy base model applied to the whole Australian population

	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
Maximum value for the ten simulations	$11,459	$52,325	30.2%	36.6%
Minimum value for the ten simulations	$11,320	$51,914	29.8%	36.4%
Difference (maximum - minimum)	$139	$411	0.5%	0.3%

Table 6.8: Results for sigmoidoscopy screening from ages 45 to 85 at different screening intervals

Screening interval	Cost per YLS	Cost per death averted	% fall in YLL due to screening	% fall in deaths due to screening
1 years	$11,564	$52,988	30.2%	36.6%
2 years	$6,667	$30,318	29.9%	36.6%
3 years	$4,862	$22,040	29.1%	35.7%
4 years	$4,001	$18,126	28.0%	34.5%
5 years	$3,539	$15,962	26.7%	33.1%
6 years	$3,283	$14,800	25.8%	31.9%

Table 6.9: Results for sigmoidoscopy screening at three year intervals for different age ranges

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
ages 45-85	$4,862	$22,040	29.1%	35.7%
ages 40-85	$5,509	$25,930	32.1%	38.2%
ages 50-85	$4,377	$18,923	24.7%	31.9%
ages 35-85	$6,359	$30,273	33.9%	39.7%
ages 40-80	$4,658	$21,515	28.8%	34.7%
ages 45 and over	$5,080	$22,817	29.3%	36.3%

Table 6.10: Results from ten simulations of screening every three years between the ages of 40 and 85 for the synthetic cohort aged 40 at the start of screening

	% fall in deaths due to screening	% fall in YLL due to screening
Mean	69.0%	69.0%
Highest value	69.3%	69.6%
Lowest value	68.7%	68.5%
Difference (highest - lowest)	0.5%	1.1%

Table 6.11: Simulated falls in mortality due to sigmoidoscopy screening every three years up to age 85, for people taking part in the screening program starting at selected ages

Age at start of screening	% fall in deaths due to screening	% fall in YLL due to screening
40	68.9%	69.0%
70	49.7%	44.5%

Table 6.12: Average increase in years of life lived for each true positive screen, for sigmoidoscopy screening every three years up to age 85 starting at selected ages.

Age at start of screening	Average increase in years of life lived
40	2.1
70	1.9

Table 6.13: Results of sensitivity analyses for colonoscopy and sigmoidoscopy screening

Model parameter	Cost per YLS	% fall in YLL due to screening
Colonoscopy screening		
Colonoscopy every 10 years, ages 45 to 85 from standard model	$15,585	37.6%
Cost of colonoscopy increased by 50%	$25,455	N/A
Cost of colonoscopy decreased by 50%	$5,224	N/A
Same participation rates as for FOBT screening	$15,517	48.1%
Sigmoidoscopy screening		
Sigmoidoscopy every 10 years, ages 45 to 85 from standard model	$5,509	32.1%
Cost of Sigmoidoscopy increased by 50%	$8,707	N/A
Cost of Sigmoidoscopy decreased by 50%	$2,011	N/A
Same participation rates as for FOBT screening	$5,423	40.6%

6.6 Figures

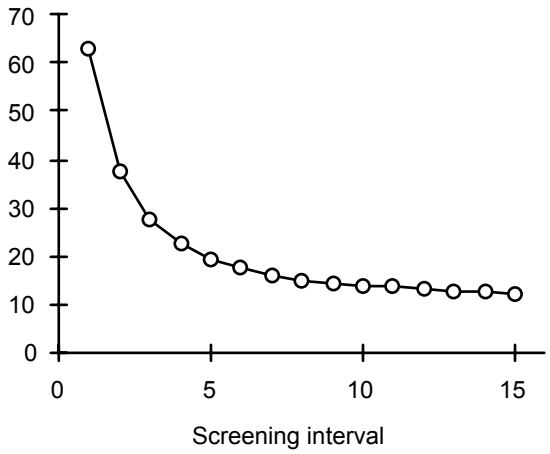

Figure 6.1: Average cost per year of life saved for population colonoscopy screening between the ages of 50 and 84 at screening intervals from one to 15 years

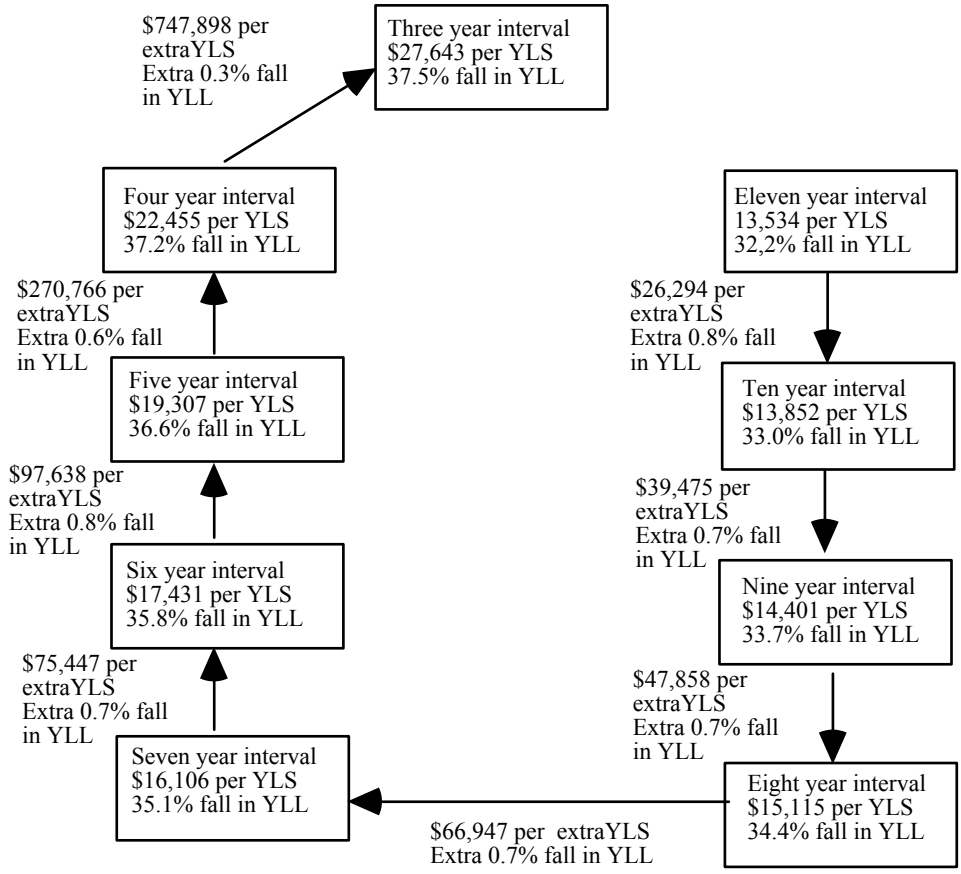

Figure 6.2: The proportion of life years saved and the extra cost per life year in moving between screening intervals for population colonoscopy screening between the ages of 50 and 84

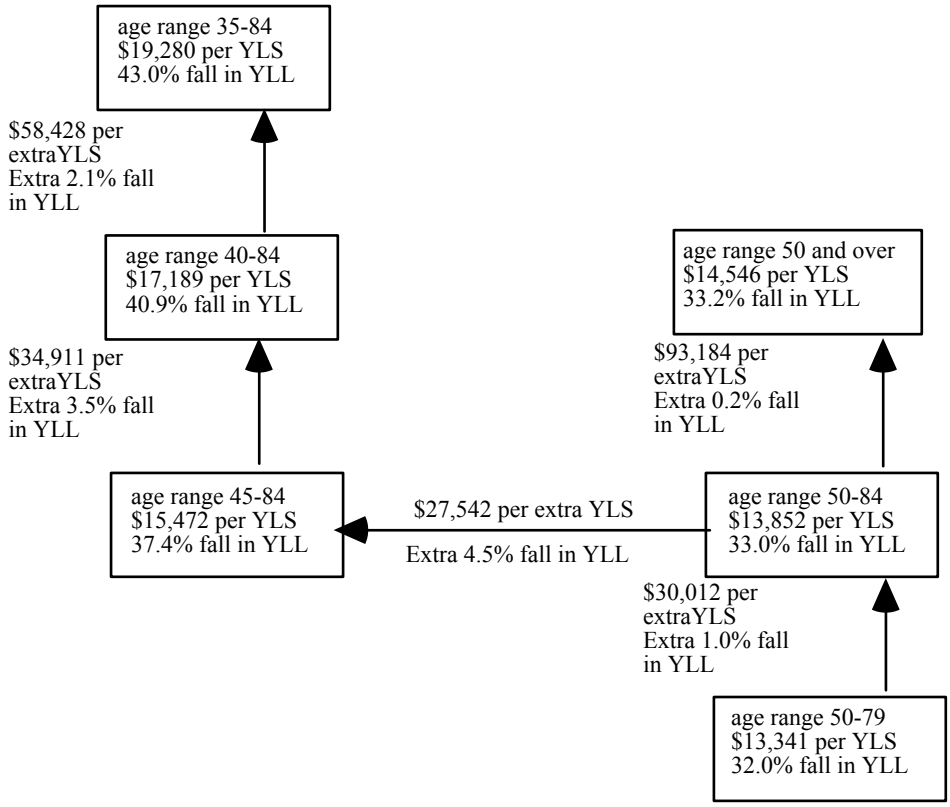

Figure 6.3: The proportion of life years saved and the extra cost per life year in moving between screening age ranges for population colonoscopy screening every ten years

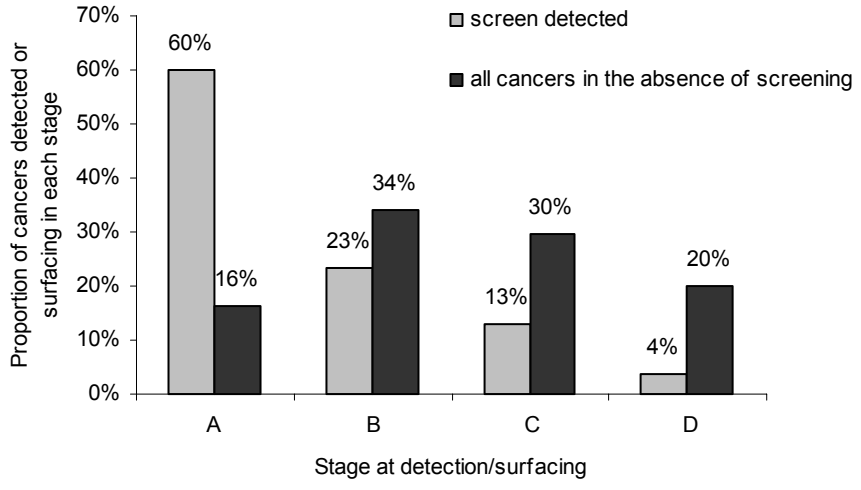

Figure 6.4: Simulated distribution of cancers by stage at detection or clinical surfacing for screening of the Australian population between ages 45 and 85 with colonoscopy every ten years

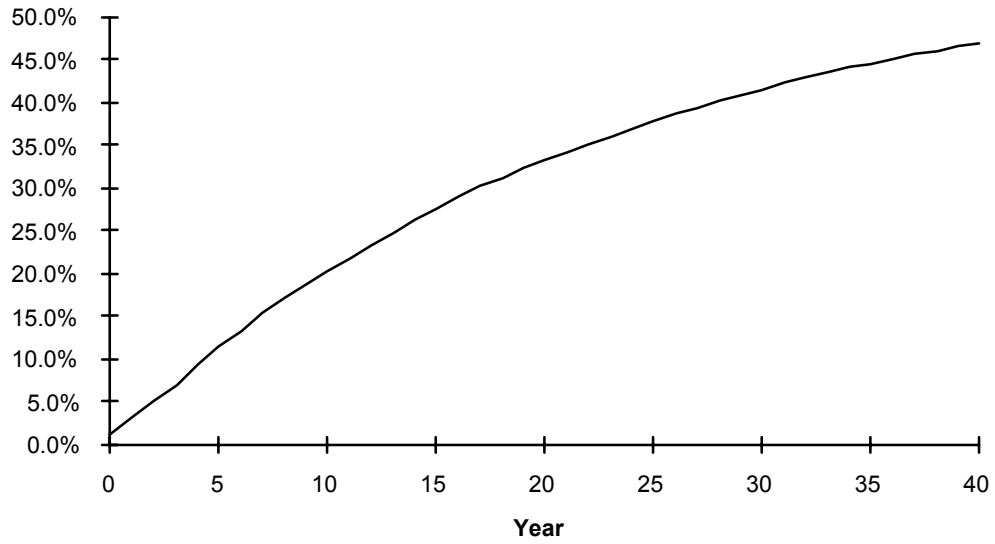

Figure 6.5: Simulated cumulative colorectal cancer mortality fall by year for screening of the Australian population between ages 45 and 85 with colonoscopy every ten years

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

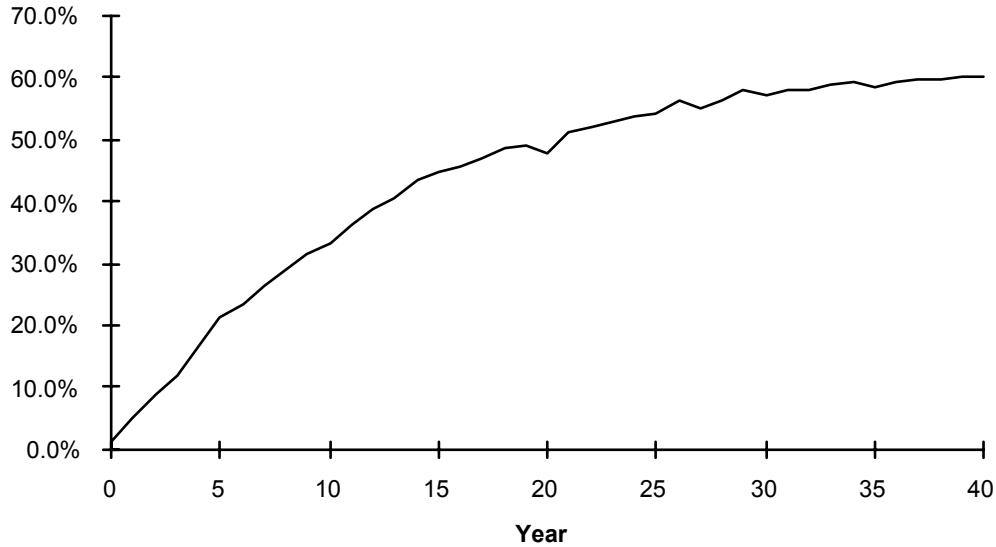

Figure 6.6: Simulated annual colorectal cancer mortality fall by year for screening of the Australian population between ages 45 and 85 with colonoscopy every ten years

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

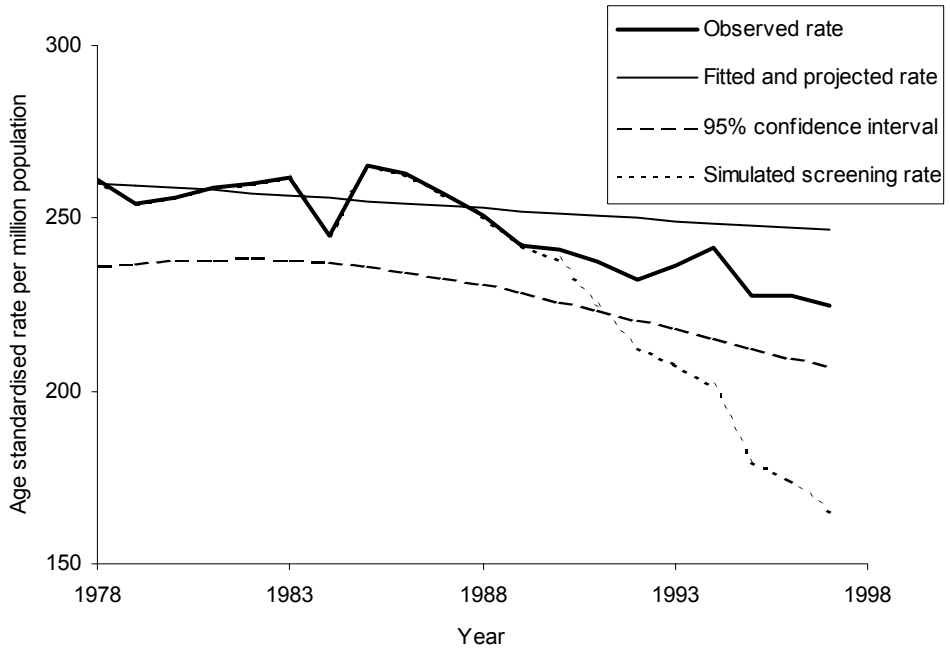

Figure 6.7: Comparison of projected colorectal cancer mortality without screening and simulated mortality with colonoscopy screening

Notes: 1. All mortality rates are age standardised to the total 1988 Australian population and calculated per million population.
2. The fitted and projected rates are calculated from a poisson regression model fitted to mortality rates from 1972 to 1989 inclusive.
3. The screened rates are calculated by applying the simulated mortality fall due to screening to the actual mortality rates for the period 1990 to 1997.

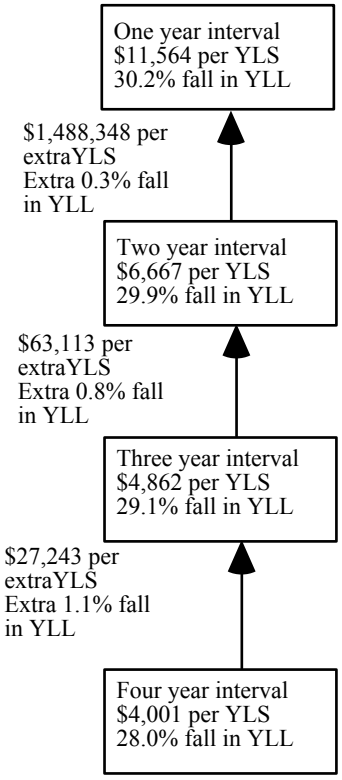

Figure 6.8: The proportion of life years saved and the extra cost per life year in moving between screening intervals for population sigmoidoscopy screening between the ages of 45 and 85

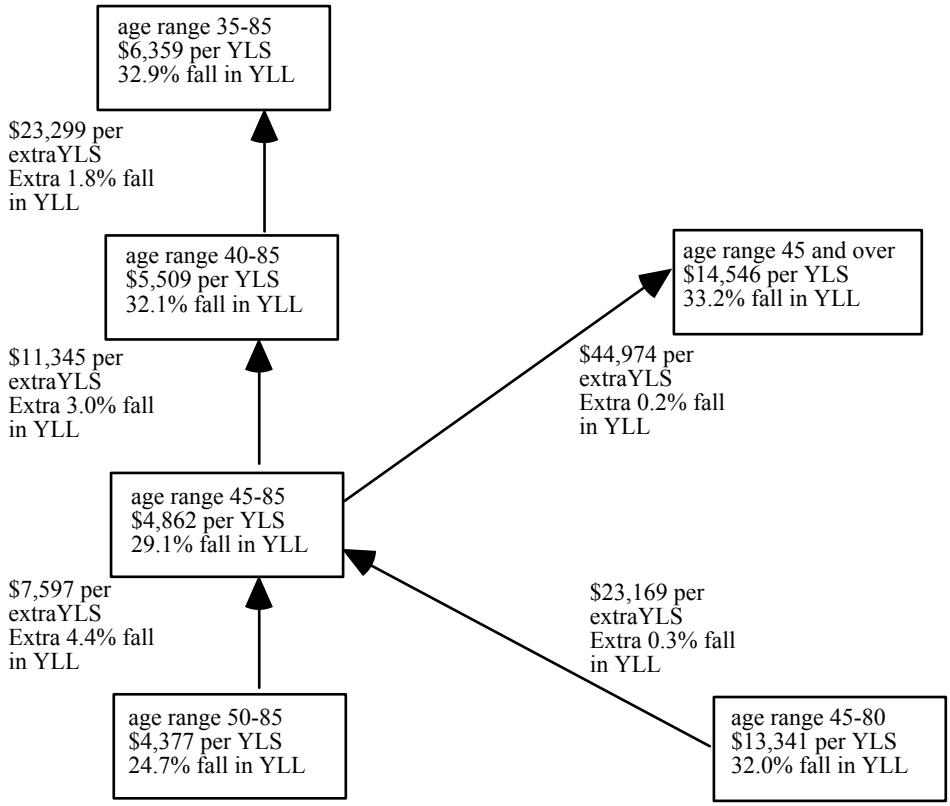

Figure 6.9: The proportion of life years saved and the extra cost per life year in moving between screening age ranges for population sigmoidoscopy screening every three years

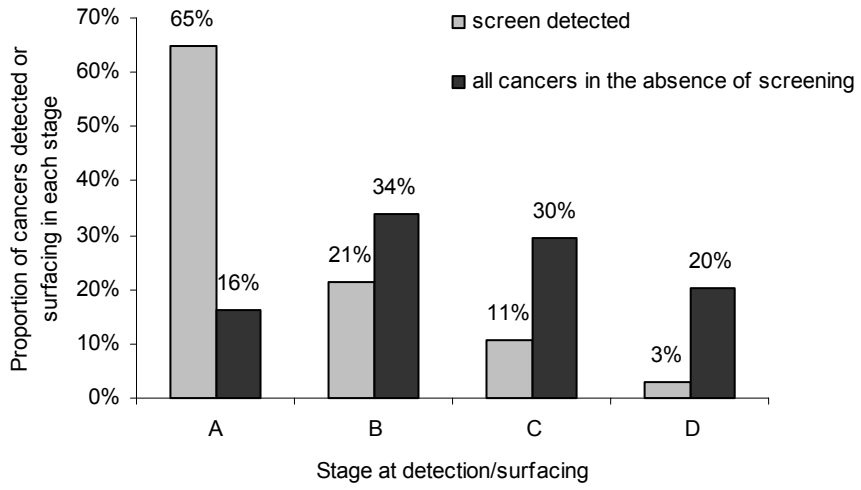

Figure 6.10: Simulated distribution of cancers by stage at detection or clinical surfacing for screening of the Australian population between ages 40 and 85 with sigmoidoscopy every three years

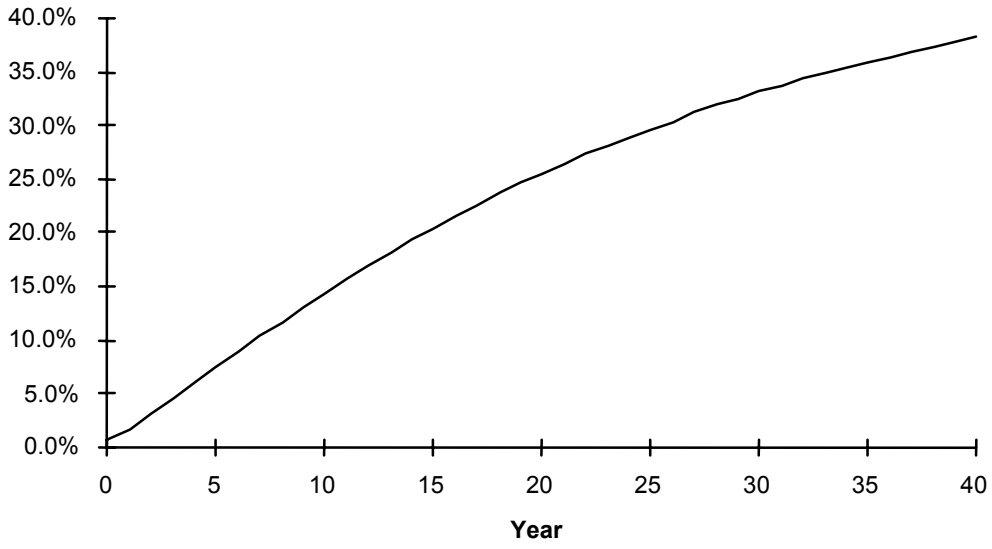

Figure 6.11: Simulated cumulative colorectal cancer mortality fall by year for screening of the Australian population between ages 40 and 85 with sigmoidoscopy every three years

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

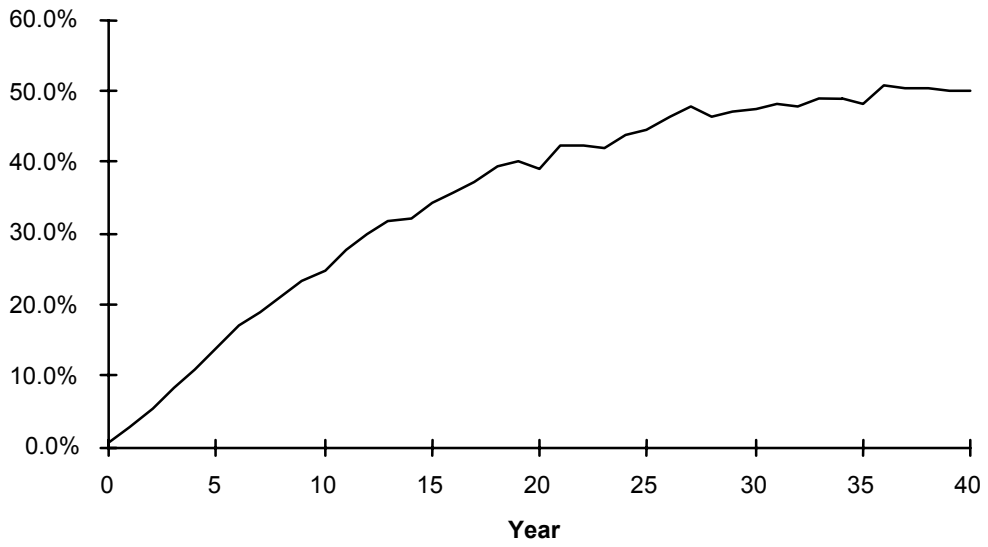

Figure 6.12: Simulated annual colorectal cancer mortality fall by year for screening of the Australian population between ages 40 and 85 with sigmoidoscopy every three years

Note: These data include all deaths from colorectal cancer irrespective of whether or not the person was in the screening program.

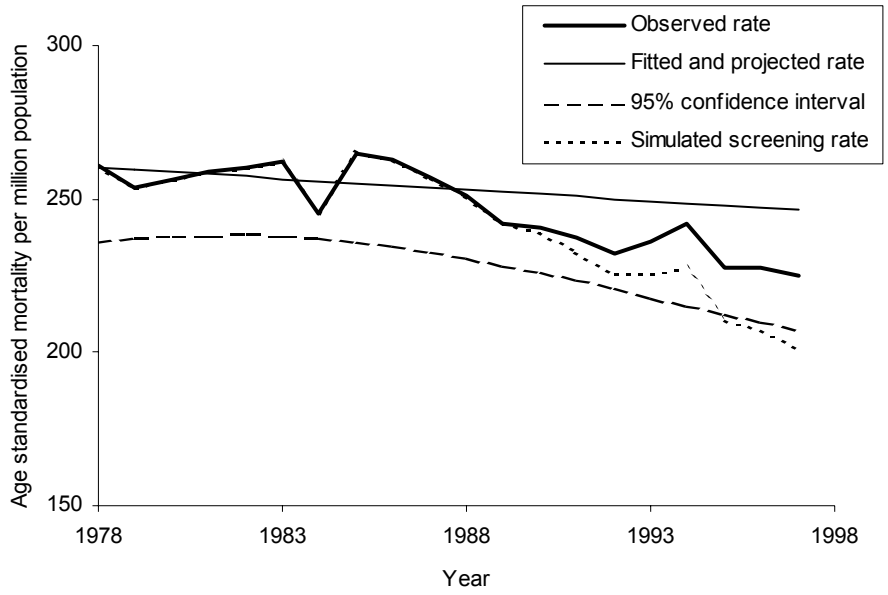

Figure 6.13: Comparison of projected colorectal cancer mortality without screening and simulated mortality with sigmoidoscopy screening

Notes: 1. All mortality rates are age standardised to the total 1988 Australian population and calculated per million population.
2. The fitted and projected rates are calculated from a poisson regression model fitted to mortality rates from 1972 to 1989 inclusive.
3. The screened rates are calculated by applying the simulated mortality fall due to screening to the actual mortality rates for the period 1990 to 1997.

7 Targeting people at high risk of colorectal cancer

7.1 Introduction

Risk factors for colorectal cancer were discussed in chapter 3. About 25% of all new cases of colorectal cancer occur in people who are at higher than average risk of the disease. While some of these cases are in people with a clearly identified genetic syndrome, who would be identified and managed outside a screening programme, the majority occur in people with a family history of colorectal cancer but without any apparent defined genetic syndrome. This chapter uses the microsimulation model to investigate specifically targeting some of these risk groups as part of a population screening programme.

The risk is moderately higher in those with one affected first degree relative and considerable higher in those with two or more affected first degree relatives or those where the relative was diagnosed at a young age (St. John & McDermott et al. 1993). Many authors consider that screening programmes should be designed to take account of this higher risk, either by starting screening at a younger age or by screening more intensively (Eddy et al. 1987, Winawer et al. 1997).

The NHMRC guidelines define categories of risk as (1) those at or slightly above average risk, (2) those at moderately increased risk and (3) those at potentially high risk (NHMRC 1999). The first group contains most of the population, including those with one affected first degree relative diagnosed at age 55 years or older. The second group contains those with an affected first degree relative diagnosed below the age of 55 or two first or second degree relatives on the same side of the family diagnosed at any age. The third group contains those with characteristics which lead to the suspicion of one of the genetic syndromes. For example it contains those with three or more first or second degree relatives on the same side of the family diagnosed with colorectal cancer—these people would be suspected to have hereditary non-polyposis colorectal cancer (HNPCC). The guidelines then suggest that all of group one be offered annual FOBT screening from the age of 50 while group two be offered colonoscopy every five years starting from the age of 50, or at age 10 years younger than the age of first diagnosis of colorectal cancer in the family, whichever comes first. They suggest that members of group 3 be subject to further diagnostic testing and, if the presence of a specific genetic syndrome is confirmed, they be managed outside a screening programme.

I will define two risk groups. The first, designated the moderate risk group, consists of those people with at least one affected first degree relative. These are the people in the first NHMRC group with the raised colorectal cancer risk and all the people in the second NHMRC group. The second group, designated the high risk group, is

a subset of the moderate risk group and consists of those people with an affected first degree relative diagnosed below the age of 55 or two first or second degree relatives on the same side of the family diagnosed at any age. This corresponds with the second NHMRC group only. Because the third NHMRC group would be managed outside a screening programme, I will not include it in my analysis.

I will simulate a screening programme where people in a risk group are offered screening by colonoscopy every five years while those in the rest of the population are offered annual FOBT screening with colonoscopy follow up of positive results. People in the risk group would be identified on the basis of a family history of colorectal cancer.

There are a number of examples in the literature of how such targeting of a risk group could be done. People with a family history of colorectal cancer could be contacted directly, by asking colorectal cancer patients to supply names and addresses of relatives. Marcus et al. (1999) have demonstrated that 91% to 96% of colorectal cancer patients agree to provide such details. Alternatively the risk group could be identified through their contact with medical services. Martyres et al. (1999) found, in a study of Victorian general practitioners, that around 97% of GPs would support an organised screening programme through their practice. Further, 95% of GPs said that they would offer screening to a patient with a family history of colorectal cancer as part of their current practice. Patients responding to an invitation to join an average risk screening programme could also be questioned on their family history and, where appropriate, assigned to a risk group. Niv (1992) provides an example of identification of a risk group by the use of a questionnaire given to people responding to an invitation to join a screening programme.

I will concentrate on screening people aged between 50 and 84. These were the age ranges identified for FOBT screening in chapter 4. In principle, the risk groups should be offered screening at ages under 50 where a relative was diagnosed with colorectal cancer at an age under 60, but there are insufficient data to model the population distribution of the risk groups below this age. Further, there are only a small number of cancers arising below age 50 which are not associated with one of the specific genetic syndromes. Hence, the exclusion of screening at ages below 50 should not greatly affect the overall results. I will run the simulation with both my medium and high risk groups as the targeted risk group and compare the results.

This simulation relies on the assumption that membership of either of my risk groups effects only the chances of developing cancer and not the course of the cancer once it has developed. This assumption is made by other modelling studies of colorectal cancer screening for people at higher risk (e.g. Eddy et al. 1987). It is also supported by Morson's observation that patients with familial adenomatous polyposis provide a good model for studying polyp growth rates which is applicable to colorectal cancer in general (Morson 1974).

7.2 Modelling parameters

In order to model a screening programme targeted at people with a family history of colorectal cancer, I need to know the proportion of people in the general population who have such a family history and the proportion of cancers which occur in people with the family history.

There are no reliable estimates of the proportion of the general population in my moderate risk group, though Eddy et al. (1987) suggest (without providing a justification) that it is about 10% for adults. Fuchs et al. (1994) in a study of nurses and health professionals found 9.4% of men and 10% of women had a history of colorectal cancer in a first degree relative. However, given the link between colorectal cancer and lifestyle and given the highly specific nature of this sample, it would be unsafe to generalise this result directly to the general population. I can get an approximate estimate using the proportion of cancers occuring in my risk group and the group's relative risk of cancer.

Both Eddy et al. (1987) and Winawer et al. (1997) say that 25% of new cases of colorectal cancer occur in people with a family history of colorectal cancer. This group includes those with a specific genetic syndrome, but Winawer et al. further break this up into HNPCC (5%), familial adenomatous polyposis (FAP) (1%) and a collection of uncommon conditions such as inflammatory bowel disease (IBD) (1%). The remainder (about 18%) they allocate to colorectal cancer family history. (Winawer et al. figure 13, page 611).

FAP, IBD and most HNPCC occur in people younger than my screening group. Further the screening protocol specifies an initial assessment of each person's risk status. This would enable people with one of these syndromes to be identified managed outside the screening programme. If I exclude all of these cancers then people with a family history account for about 19% of the remaining cancers. Hence if I assume negligible numbers of people with a specific genetic syndrome in the screening group, then this is the proportion of cancers in the screening group which occur in the medium risk group.

St. John & McDermott et al. (1993) derived an odds ratio for colorectal cancer in people with a family history at ages above 45 of 1.8. Their analyses also showed no detectable difference in odds ratio according to gender, site of cancer or type of cancer (single versus multiple). I use this value as my estimate of relative risk. If I let

α_h = the proportion of all cancers which occur in the risk group;

p_h = the prevalence of the risk factor in the screening target group;

RR = the risk of cancer in the risk group, relative to people not in the group;

r_h = the risk of cancer in the risk group; and

r_a = the risk of cancer in the target population not in the risk group.

Then

$$\alpha_h = p_h r_h \Big/ \left[(1-p_h)r_a + p_h r_h\right] = p_h RR \Big/ \left(1 - p_h + p_h RR\right) \quad (1)$$

Rearranging the terms in this equation gives me

$$p_h = \alpha_h \Big/ \left[\alpha_h + RR(1-\alpha_h)\right] \quad (2)$$

This equation gives me a population prevalence for my moderate risk group of 12%.

The NHMRC guidelines state that my high risk group comprises one to two per cent of the population and has a three to six fold increase in risk (NHMRC 1999). St. John & McDermott et al. suggest that the relative risk when the family history is at ages below 45 is 3.7 while the relative risk for a family history of two or more cases is around 6. This suggests that the combined relative risk for my high risk group is between 3 and 6, so I will take the mid point relative risk of 4.5 and investigate the effect of varying this in the sensitivity analysis. If I assume the higher population prevalence of 2%, equation (1) gives me an estimate of the proportion of cancers occuring in the high risk group as 8%.

Niv (1992) investigated the use of a questionnaire to identify people at higher risk and found that of 633 people identified as having risk factors, 31% underwent colonoscopy. Niv's risk group is broader than either of my risk groups (for example, people were asked about first degree relatives having cancers other than colon cancer), but it is close to my medium risk group. It is not clear from Niv's paper whether people chose colonoscopy screening or were allocated to it (or to some other screening mode). However, the result shows that at least 31% colonoscopy screening among the medium risk group is an achievable result. Colombo et al. (1997) found a similar participation rate (30%) in first degree relatives of colorectal cancer patients who were offered sigmoidoscopy screening. Marcus et al. (1999) note that screening programmes targeted at my medium risk group have participation rates that rarely exceed 40% to 50%. I will use Niv's result of 31% as the acceptance rate for colonoscopy screening in my medium risk group and investigate higher screening participation as part of the sensitivity analyses.

This screening programme includes annual FOBT screening for people at average risk and it is likely that some of the risk group who refuse a relatively invasive screening test such as colonoscopy will instead choose FOBT screening. Niv found that a total of 78% of people identified as at risk using his questionnaire accepted some form of screening (including partial colonoscopy, barium enema and FOBT testing). I will use 78% as the total screening participation rate for my medium risk group, so that 31% accept colonoscopy and the remaining 47% join the annual FOBT screening programme.

Houlston et al. (1990) describe a colonoscopy screening programme offered to people in my high risk group which attained a participation rate of 90%. It seems logical that most people identified at high risk will accept screening, so I will accept this as my participation rate for the high risk group. In principle, the remainder may opt for the annual FOBT screening programme. However, varying the FOBT screening rate for this colonoscopy screening refusal group from 0% to 100% made

negligible difference to the results, so I will arbitrarily set the FOBT screening participation rate for those who refuse colonoscopy to zero.

The model parameters for targeted screening are listed in table 7.1.

The synthetic population was generated by randomly assigning population members with cancer to be either in or out of the risk group with a probability equal to the proportion listed in table 7.1. Similarly, members without cancer were randomly assigned to the risk group so that the overall prevalence was equal to the population prevalence of the risk group listed in the table.

7.3 Results

Table 7.2 lists the percentage fall in cumulative mortality and cumulative discounted life years lost and the cost per life year saved and per death averted for ten simulations of the medium risk model applied to the whole Australian population. There is less than half a per cent difference between the highest and lowest values of the in cumulative deaths prevented and YLS. The highest and lowest cost per life year saved differ by less than $150 while the highest and lowest cost per death prevented differ by around $700. The differences between screening protocols in each of these measures would have to be greater than this to have any practical implications.

Table 7.3 lists the saving in deaths and YLL and the cost per death averted and per YLS for annual FOBT screening for people at average risk and colonoscopy screening targeting the medium or high risk group. Table 7.3 also includes the equivalent results for annual FOBT screening without targeting, taken from chapter 5, for comparison. Figure 7.1 presents the marginal costs and benefits of moving from no targeting to targeting the medium and high risk groups.

The savings in mortality and the average cost per YLL and per death averted in the programmes which include targeted screening for the medium and high risk groups are very similar. If I look at the extra costs and benefits relative to the base screening programme, the extra cost of including targeted screening for the high risk group is $12,274 per YLS which produces an extra 1.6% fall in YLL. The cost of including targeted screening for the medium risk group is $20,535 per YLS which produces an extra 2.7% fall in YLL. The extra cost of moving from targeting high risk only to targeting the medium risk group is $32,758 which produces and extra 1% fall in YLL. This is close to my $32,000 limit at which governments will fund health interventions on the basis of cost-effectiveness. However, that limit was only approximate and including colonoscopy screening for the medium risk group leads to an extra savings in YLL around two thirds as big again as the saving from only targeting the high risk group.

These results support targeting people in the high risk group with more intensive screening, as suggested by the NHMRC guidelines. However, they also suggest that the NHMRC recommendation of only targeting the high risk group for colonoscopy may be too conservative. The identification of people in the medium risk group and

offering them colonoscopy screening may be a cost-effective alternative to the programme which only offers annual FOBT screening.

7.4 Sensitivity analyses

Three sensitivity analyses were done for the medium risk group. The first set the prevalence of the medium risk group in the general population to the value of 10% suggested by Eddy et al. (1987) and Fuchs et al. (1994). The second set the colonoscopy screening rate at 50%, which is the highest value suggested by Marcus et al. (1999). The third was based on the assumption that all risk group members who refused colonoscopy would join the annual FOBT screening programme. Table 7.4 lists the average cost per YLL and per death averted and the savings in YLL and deaths under each alternative scenario. All scenarios lead to similar average savings in YLL and deaths. However, the interesting question is whether each scenario would lead to the same decision as to whether or not to offer colonoscopy to people in the medium risk group.

The extra cost per YLS and the extra gain in YLL in moving from targeting only the high risk group with colonoscopy screening to targeting the medium risk group under each alternative scenario are listed in table 7.5. If the population prevalence of the medium risk group is 10% rather than 12%, then the choice between targeting the high and medium risk groups is more marginal. The cost per extra YLS is $36,062, which is higher than my limit of $32,000 but not by so much as to rule out targeting the medium risk group. However, the extra gains in targeting the medium risk group are much smaller. Achieving a colonoscopy participation rate of 50% among the medium risk group does not substantially alter the choice between targeting the medium or the high risk group. Finally, if all the medium risk group participate in screening (either colonoscopy or FOBT), then the choice to offer colonoscopy to the medium risk group is clearly a cost-effective one.

Five sensitivity analyses were done for the high risk group. Two analyses involved setting the relative risk of cancer in the high risk group at 3 and at 6, in line with the range suggested by the NHMRC guidelines report (NHMRC 1999). The third involved setting the population prevalence of the high risk group to 1%—the lower value suggested by the NHMRC report. The fourth analysis assumed that participation in colonoscopy screening was only 50% (the highest value suggested by Marcus et al., 1999) rather than 90%. The final analysis assumed that all risk group members who refused colonoscopy would join the annual FOBT screening programme.

Table 7.4 also lists the average cost per YLL and per death averted and the savings in YLL and deaths under each alternative scenario for the high risk group. Again all scenarios lead to similar average savings in YLL and deaths. The extra cost per YLS and the extra gain in YLL in moving from targeting only the high risk group with colonoscopy screening to targeting the medium risk group under each alternative scenario for the high risk group are listed in table 7.6.

Cost-effectiveness considerations would lead to targeting the medium risk group rather than just the high risk group under each alternate scenario except for the case when the high risk group has a six-fold increase in the risk of developing colorectal cancer. In this case, the cost per extra YLS in offering colonoscopy to the medium risk group rather than only the high risk group is $70,016, leading to an extra 0.6% gain in YLS, which is well above my $32,000 limit and for a relatively small gain in YLS. This suggests that if the high risk group has a relative risk of 6 or greater, then the gains in offering colonoscopy screening to the whole medium risk group would not be worth the extra costs.

7.5 Tables

Table 7.1: Model parameters for screening people in medium and high risk groups

Medium risk group	
Proportion of all cancers in the screening target group occuring in the medium risk group[1]	0.19
Relative risk for colorectal cancer in people in the medium risk group[2]	1.8
Proportion of the screening target group in the medium risk group	0.12
Colonoscopy participation rate for people in the medium risk group[3]	0.31
Total screening participation rate for people in the medium risk group[3]	0.78
High risk group	
Proportion of all cancers in the screening target group occuring in the high risk group	0.08
Relative risk for colorectal cancer in people in the high risk group[4]	4.5
Proportion of the screening target group in the high risk group[4]	0.02
Screening participation rate for people in the high risk group[5]	0.90

Sources:
1. Derived from data presented in Winawer et al. (1997)
2. St. John & McDermott et al. (1993)
3. Niv (1992)
4. NHMRC guidelines report (NHMRC 1999)
5. Houlston et al. (1990).

Table 7.2: Results from ten simulations of the medium risk model applied to the whole Australian population

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Maximum value for the ten simulations	$10,082	$47,475	31.5%	37.2%
Minimum value for the ten simulations	$9,938	$46,776	31.0%	36.8%
Difference (maximum - minimum)	$144	$699	0.4%	0.4%

Table 7.3: Results for targeting medium and high risk groups

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Annual FOBT at ages 50 to 84	$8,987	$41,777	28.5%	34.2%
Screening ages 50 to 84, 5 yearly colonoscopy for medium risk group, annual FOBT for remaining population	$10,000	$47,230	31.2%	36.9%
Screening ages 50 to 84, 5 yearly colonoscopy for high risk group, annual FOBT for remaining population	$9,165	$43,206	30.1%	35.6%

Table 7.4: Sensitivity analyses

	Cost per YLS	Cost per death	% fall in YLL due to screening	% fall in deaths due to screening
Medium risk group				
Base model	$10,000	$47,230	31.2%	36.9%
Risk group comprises 10% of general population	$9,798	$45,969	30.7%	36.5%
Risk group achieves 50% participation rate in colonoscopy screening	$10,527	$49,729	32.0%	37.8%
All members of risk group who refuse colonoscopy screening join the annual FOBT screening program	$9,727	$46,287	31.5%	37.0%
High risk group				
Base model	$9,165	$43,206	30.1%	35.6%
Risk group has relative risk of 3	$9,213	$43,517	29.9%	35.4%
Risk group has relative risk of 6	$8,973	$42,540	30.6%	36.1%
Risk group comprises 1% of general population	$8,942	$42,107	29.6%	35.0%
Risk group only achieves 50% participation rate in colonoscopy screening	$9,210	$43,035	29.3%	34.9%
All members of risk group who refuse colonoscopy screening join the annual FOBT screening program	$9,158	$43,160	30.2%	35.6%

Table 7.5: The marginal gains in moving from offering colonoscopy screening only to the high risk group to offering colonoscopy to the medium risk group under each alternative scenario for the medium risk group

	Cost per extra YLS	Cost per extra death averted	extra % fall in YLL due to screening	extra % fall in deaths due to screening
Risk group comprises 10% of general population	$36,062	$141,185	0.6%	0.9%
Risk group achieves 50% participation rate in colonoscopy screening	$32,535	$158,989	1.8%	2.2%
All members of risk group who refuse colonoscopy screening join the annual FOBT screening program	$21,551	$128,007	1.3%	1.3%

Table 7.6: The marginal gains in moving from offering colonoscopy screening only to the high risk group to offering colonoscopy to the medium risk group under each alternative scenario for the high risk group

	Cost per extra YLS	Cost per extra death averted	extra % fall in YLL due to screening	extra % fall in deaths due to screening
Risk group has relative risk of 3	$28,991	$136,444	1.3%	1.5%
Risk group has relative risk of 6	$70,016	$270,457	0.6%	0.8%
Risk group comprises 1% of general population	$30,193	$151,196	1.6%	1.8%
Risk group only achieves 50% participation rate in colonoscopy screening	$21,874	$123,324	1.9%	1.9%
All members of risk group who refuse colonoscopy screening join the annual FOBT screening program	$32,687	$163,932	1.0%	1.2%

7.6 Figure

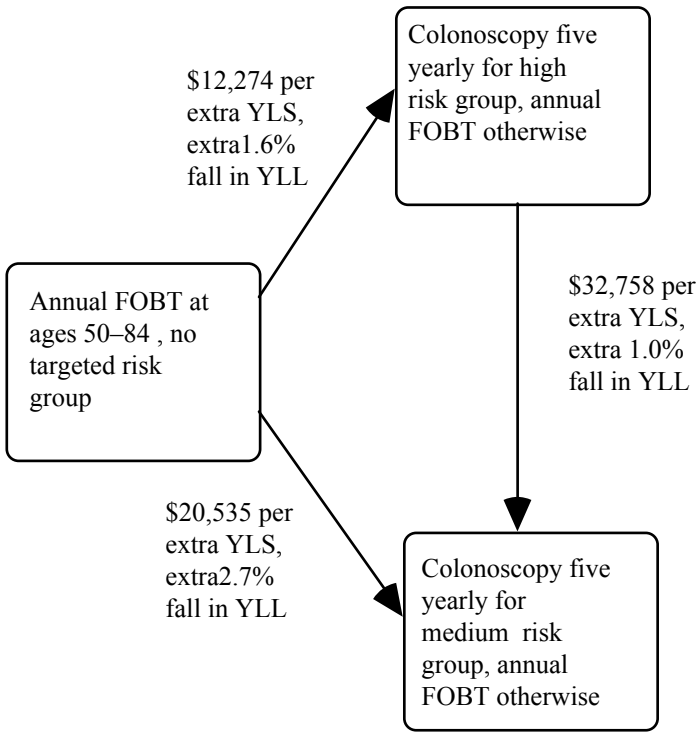

Figure 7.1: Change in YLL and cost per extra life year saved when moving between no targeted risk group and targeting medium and high risk groups

8 Conclusions

At the start of this thesis I posed three questions which this study set out to answer. These were:

- What are the likely benefits and costs of colorectal cancer screening in Australia;
- How should such screening be carried out; and
- What more do we need to know to design cost-effective screening programmes?

In this chapter I present the costs and benefits of a population screening programme based on each of FOBT, colonoscopy and sigmoidoscopy. I also describe the most appropriate screening protocol for such a programme for each of the screening modalities. Finally I present some future directions for study in this area and propose priorities for addressing gaps in our knowledge of screening for colorectal cancer.

8.1 A population FOBT screening programme

8.1.1 Standard guaiac FOBT screening using unrehydrated slides

My results support the NHMRC recommendations that asymptomatic people with no prior indications of colorectal cancer risk be offered annual FOBT screening commencing at age 50. However, I found that an organised population screening programme that continued screening after age 84 could not be justified on cost-effectiveness grounds. That is not to say that such screening should not be done. Individuals may decide that the small potential mortality gains are worth the extra personal expense. However, these mortality gains do not justify the diversion of public funds from other health areas to fund screening at ages over 84. On the other hand the cost-effectiveness of a publicly funded organised programme of annual FOBT screening, with colonoscopy follow-up of positive results, targeting people aged 50 to 84 is consistent both with other government funded health interventions and with the cost per YLS which governments appear willing to pay.

My simulated FOBT screening programme had a cost per YLS of $8,987 and a cost per death averted of $41,777. The simulated proportional fall in YLL due to colorectal cancer was 28.5% while the simulated proportional fall in deaths from colorectal cancer was 34.2%. These results apply to a screening period of 40 years. This is important in their interpretation. As we saw when validating my model

against the published results from the RCT's, a shorter screening period can give rise to a lower cumulative mortality gain with a higher cost per YLS.

I found that there were significant benefits to participants in the screening programme. I looked at benefits to people who joined the programme at age 50 and to people who didn't join until age 70. Obviously the larger gains come from starting screening at age 50. The average risk of death from colorectal cancer for screening participants over the rest of their life fell by 68%, while the average years of life lost due to colorectal cancer fell by 62%. However, the gains from screening remain substantial even for those who enter the screening programme for the first time at age 70. For these people the average risk of death from colorectal cancer over the rest of their life fell by 52%, while the average years of life lost due to colorectal cancer fell by 45%. Each positive screen resulted in an average of 2.1 extra years of life for people starting screening at age 50 and an extra 1.7 years of life for people starting screening at age 70.

I found that by using standard analysis of currently available national mortality data, the first mortality benefits from a national screening programme should be detectable after around three years of screening. Most of the decrease in annual colorectal cancer mortality will be achieved after around 14 years, though this annual rate will continue falling slowly until at least 30 years. It is important to keep this in mind when designing a monitoring system for a national screening programme. Most public health policy is developed with a much shorter time frame than 30 or even 14 years and it is important that monitoring and evaluation of screening programmes is not based on an expectation of large short-term mortality gains.

8.1.2 Other FOBT screening programmes

I looked at rehydration of FOBT slides to increase screening sensitivity. This rehydration also decreases the screening specificity and hence increases the screening cost. My results support the practice of the most recent RCTs in using unrehydrated slides. I found that the small mortality gains in using rehydrated slides were more than offset by the large extra cost per YLS arising from the higher rate of false positive screen results.

I also looked at the use of immunochemical FOBTs. These have not been studied as part of a population screening programme, so my modelling parameter values were only approximate. However, my results suggest that immunochemical FOBT screening could provide a cost-effective alternative to the standard guaiac FOBT in a population screening programme. The simulated cost per YLS using immunochemical FOBT was around $1,000 lower than the cost per YLS for the guaiac FOBT and the immunochemical FOBT screening programme led to an extra 10% YLS.

Finally I looked at combining FOBT and sigmoidoscopy screening. There have been no major studies of combined FOBT and sigmoidoscopy screening, so there are no data on which to base estimates of screening participation rates. However I derived

a plausible range of screening participation rates and simulated a range of costs and benefits which could be expected from such screening. I simulated combining sigmoidoscopy and FOBT screening by offering people sigmoidoscopy screening every 5 years with annual FOBT screening in the intervening years. Such a programme with a lower participation rate led to a cost per YLS which was almost equal to FOBT screening without sigmoidoscopy. If the higher participation rate was realised then the cost per YLS fell by around $2,000. The extra YLS compared to FOBT screening without sigmoidoscopy ranged from around 2.6% for the lower participation rate to almost 11% for the higher participation rate. These results suggest that the addition of sigmoidoscopy every 5 years to the annual FOBT screening programme would be a cost-effective way of increasing the screening mortality gains.

I also looked at the benefits to participants in the combined screening programme. I found that the addition of 5 yearly sigmoidoscopy to an annual FOBT screening programme results in the prevention of an extra 14% deaths from colorectal cancer, leading to an extra 15% fall in YLL.

8.2 Population screening programmes based on colonoscopy and sigmoidoscopy

Both Winawer et al. and the NHMRC suggest that flexible sigmoidoscopy is an acceptable screening tool if offered every 5 years at ages 50 and over. Winawer et al. also suggest screening at ages 50 and over with colonoscopy every 10 years, though they note that there are no studies evaluating screening colonoscopy alone. I examined the use of each of these modalities as the primary screening tool in a population based screening programme.

Colonoscopy

My results suggest that colonoscopy screening every 10 years starting at age 45 and finishing at age 85 would be an acceptable mass screening programme in cost-effectiveness terms. My simulated colonoscopy screening programme had a cost per YLS of $15,585 and a cost per death averted of $69,246. The simulated proportional fall in YLL due to colorectal cancer was 37.6% while the simulated proportional fall in deaths from colorectal cancer was 47.1%. As before, these results apply to a screening period of 40 years. This supports Winawer et al.'s suggestion for colonoscopy screening though I would start screening at a younger age and, as with FOBT screening, I would include an upper age limit. My results imply that a programme offering screening over the age of 85 would not be justified on cost-effectiveness grounds.

When I simulated the benefits to screening participants I found that for a cohort aged 45 at the start of screening the average risk of death from colorectal cancer over the rest of their life fell by 79%, while the average years of life lost due to colorectal cancer fell by 76%. For the cohort which entered the screening

programme at age 70 the average risk of death from colorectal cancer over the rest of their life fell by 74%, while the average years of life lost due to colorectal cancer fell by 69%. Each positive screen resulted in an average of 2.2 extra years of life for people starting screening at age 45 and an extra 2.0 years of life for people starting screening at age 70.

As with FOBT screening, the first mortality benefits from a national screening programme should be detectable using standard analysis of currently available national mortality data after around three years of screening. The simulation also suggested that, while substantial gains can be demonstrated after 10 to 15 years of screening, the programme would not achieve its full potential gains before the end of the 40 year screening period.

Sigmoidoscopy

My results suggest that sigmoidoscopy screening every 3 years starting at age 40 and finishing at age 85 would be an acceptable mass screening programme in cost-effectiveness terms. My simulated sigmoidoscopy screening programme had a cost per YLS of $4,862 and a cost per death averted of $22,040. The simulated proportional fall in YLL due to colorectal cancer was 29.1% while the simulated proportional fall in deaths from colorectal cancer was 35.7%.

For a cohort of screening participants aged 40 at the start of screening, both the average risk of death from colorectal cancer over the rest of their life and the average years of life lost due to colorectal cancer fell by around 69%. For a cohort entering the screening programme at age 70 the average risk of death from colorectal cancer over the rest of their life fell by 50%, while the average years of life lost due to colorectal cancer fell by 45%. Each positive screen resulted in an average of 2.1 extra years of life for people starting screening at age 40 and an extra 1.9 years of life for people starting screening at age 70.

A national monitoring system should be able to demonstrate that screening was having an effect on mortality after six years of a sigmoidoscopy screening programme. However, the simulation suggests that while substantial gains can be demonstrated after around 20 years of screening with the programme achieving its full potential gains after around 35 years.

The screening interval arising from my simulation is shorter than the 5 years recommended by both sets of guidelines. The NHMRC guidelines give no justification for the screening interval. Winawer et al. cite studies that show (1) colonoscopy is equally effective at intervals of 1 and 3 years, (2) sigmoidoscopy can be effective at intervals of up to 10 years (3) few polyps arise and progress to cancer in a 5 year period. However, elsewhere in their paper they also cite studies that show

"…adenomatous polyps rarely arise and progress to cancer over periods less than 3 years." (p 608)

Further, the case-control study of sigmoidoscopy screening that Winawer et al. cite in justifying the effectiveness of such screening found that out of a total of 261 cases

with fatal cancers, there were 7 whose cancer was within the reach of sigmoidoscopy and was diagnosed from 1 to 3 years after a negative sigmoidoscopy screen and a further 6 whose cancer was diagnosed within 4 years (Selby et al. 1992). This is at least consistent with a screening interval of less than 5 years.

The cost per YLS from this simulation is much lower than that arising from any of the other simulations. This is not surprising for colonoscopy, which has a much higher cost per screen than sigmoidoscopy. However, it is unexpected for FOBT screening. This lower cost per YLS arises in two ways. The first is the fact that sigmoidoscopy screening is applied every three years compared to FOBT screening which is annual. The second is that sigmoidoscopy does not produce false positive results which need to be further investigated by colonoscopy in the way that FOBT screening does.

The benefits of sigmoidoscopy screening both to participants and in terms of overall colorectal cancer mortality are similar to FOBT screening despite the much lower initial screening participation rate. This arises from a person's ability to join the screening programme after an initial refusal. The model hypothesises a situation in which the initial refusers are offered screening every year until they either join the programme or pass the programme's upper age limit. Once in the programme, each participant has a high probability of remaining in the programme and can only exit once every three years (when actually offered screening). Hence over time the actual participation rate grows. This allows sigmoidoscopy screening to achieve an effective participation rate similar to that achieved by FOBT screening. It is unclear whether this effect would be observed in practice with a sigmoidoscopy screening programme.

These results suggest that screening with sigmoidoscopy as the primary screening tool is at least as effective as annual FOBT screening and may be cheaper. However, as the NHMRC guidelines note,

"FOBT and sigmoidoscopy are complementary in that FOBT has the potential to detect lesions proximal to the reach of the sigmoidoscope." (p 27)

and given the results from simulating the combined sigmoidoscopy and FOBT screening, it is more likely that sigmoidoscopy would be used in combination with another screening modality.

8.3 Screening people at high risk of colorectal cancer

My results support the NHMRC recommendations that people at higher risk based on a family history of colorectal cancer should be offered colonoscopy screening. However, if that offer occurs in the context of a programme offering annual FOBT screening to people at average risk, then the NHMRC guidelines may be too conservative. The guidelines suggest that the offer should only be made to those in my high risk group whereas my results suggest that offering it to people in my medium risk group may also lead to cost effective mortality savings.

I simulated an annual FOBT screening programme which included identification of people in a risk group and offering them five yearly colonoscopy as an alternative to annual FOBT. There were moderate mortality gains with a low to moderate cost per YLS for both the simulation where the risk group was my high risk group and where it was my medium risk group. When I examined the extra mortality gains and cost per YLS in including the moderate risk group compared to the high risk group, the results suggested that the moderate risk group should probably also be targeted for colonoscopy screening. However, sensitivity analysis showed that this choice was different for a range of plausible alternative model parameters.

The mortality savings for both high and medium risk groups are not large because the group with the very high risk is not large and the medium risk group which is relatively large does not have a greatly elevated risk. Further, some of the mortality gains in the medium risk group arise from the assumption that the identification of their risk will lead to greater participation in FOBT testing among those in the risk group who refuse colonoscopy. However, given the current level of knowledge of the size of the risk groups and their level of elevated risk, I would probably recommend the screening programme include identification and the offer of colonoscopy for the medium risk group rather than just the high risk group.

My simulation of a combined FOBT and colonoscopy screening programme, where colonoscopy was offered to the medium risk group, had a cost per YLS of $10,000 and a cost per death averted of $47,230. The simulated proportional fall in YLL due to colorectal cancer was 31.2% while the simulated proportional fall in deaths from colorectal cancer was 36.9%%.

8.4 Conclusions and future directions

The results of this study clearly show that a population based screening programme for colorectal cancer would be a cost-effective publicly funded health intervention. While the RCTs have focussed on FOBT screening with colonoscopy follow-up, my results suggest that a cost-effective screening programme could also be based on sigmoidoscopy or colonoscopy. However, in view of practical difficulties such as ensuring an adequate supply and geographic distribution of practitioners trained in endoscopy, a FOBT screening programme is more likely to be implemented, possibly with some colonoscopy screening targeted at people in risk groups. This section presents some priorities for future research in this area to fill some of the gaps in our knowledge of colorectal cancer screening in order to better plan such a screening programme.

Disease incidence and natural history

Data from the RCTs, when it is published, could be used to provide better estimates of some of the more uncertain of the disease model parameters such as the sojourn times in each disease stage. However, my model validation results suggest that the disease model is adequate for simulating at least FOBT screening while our sensitivity analyses suggest that my results are relatively robust to moderate

changes in the least certain disease model parameters. Hence such improved estimates would not be a high priority in planning a screening programme.

A more useful focus of future work would be to improve estimates of the proportion of the population at higher than average risk of colorectal cancer and the nature of that risk. The sensitivity analyses conducted with our simulation of screening people at higher risk showed that the results were dependent on which of several plausible values was chosen for the number of people in each of our risk groups and the size of their relative risk. Since the risk is defined in terms of a person's family history of colorectal cancer, this information should be relatively easy to collect. The medium risk group would be large enough to investigate using population surveys such as the Australian Bureau of Statistics National Health Survey, or one of the state based routine health surveys. The high risk group would be better suited to a 'case-finding' study such as that proposed by Marcus et al., where people diagnosed with colorectal cancer were asked about their first degree relatives.

Screening model parameters

My sensitivity analyses highlighted the key role of test sensitivity and specificity in the model's simulations. It is difficult to find a good source for these estimates in the context of a population screening programme. Smaller studies, where, for example, there is greater ability to ensure quality control in the storage, handling and interpretation of FOBT slides, may give overly optimistic results. Data from the larger RCTs are derived in a population screening context but do not yet provide adequate estimates. However, my model validation results suggest that the screening model is adequate for simulating at least FOBT screening. Hence the sensitivity estimates should be adequate for planning an FOBT screening programme. Similarly my model reproduced the cost per YLS proposed by Salkeld et al. As noted in chapter 4, this was not a validation of the cost estimates per se, but did show that my model's aggregation of the screening costs, which includes the specificity, is consistent with the results of the Minnesota study.

Test sensitivity and specificity are not just properties of the test alone but depend as well on the context in which the test is performed. In this case they would depend on factors such as how well good quality control could be maintained over the storage, handling and interpretation of FOBT slides in a the context of the Australian population screening programme. Hence it would be important to know that my sensitivity and specificity estimates could be achieved in practice in the Australian context. However, further refining these estimates would not be the first priority of further studies.

A more useful focus would be to derive better estimates of sensitivity and specificity of alternative FOBT screening such as immunochemical FOBT. My simulations suggest that immunochemical FOBT would be a cost-effective alternative to the standard guaiac FOBT but there are no studies to show how well this would perform in practice in a population screening programme. Better understanding of the test sensitivity and specificity, as well as any practical problems in

implementing immunochemical FOBT screening on a large scale, would be a priority in deciding which FOBT to use as the basis for a screening programme.

My sensitivity analyses also showed that the population screening participation rate is a key model parameter which should be a high priority for investigation. This would have two aspects. Firstly we have screening participation rates for England and parts of Europe from the RCTs but we need to know how achievable these would be in the Australian population. Secondly, the RCT results suggest that significant increases in participation can be achieved with more intensive recruitment. We need to know what recruitment strategies are most likely to work for the Australian population. We also need to know how well people would accept modifications to the screening programme such as targeting people at higher risk or including sigmoidoscopy in a programme for people at average risk. Experience from screening for breast and cervical cancer suggests that we also need to consider these issues for specific population subgroups such as indigenous Australians and people from a non-English speaking background.

Both breast and cervical cancer screening programmes were subject to an extensive series of pilot projects as part of developing the national programme. (AHMAC 1990, AHMAC 1991) The planning of a colorectal cancer screening programme would probably proceed in the same way. If so, the pilot projects would be the appropriate vehicle for examining issues of sensitivity, specificity and screening participation in practice.

Costs

A detailed investigation of screening costs was beyond the scope of this study, but this is an area which should be a priority. Firstly the cost estimates I used should be refined. For example, the sensitivity analyses confirmed the key role of colonoscopy costs in the overall screening costs. I followed Salkeld et al. in assuming a single cost for colonoscopy. However, colonoscopy costs will vary, depending on whether or not a polyp is found which requires removal and biopsy.

Secondly, while most of the published cost-effectiveness models for colorectal cancer have restricted costs to a similar set to mine, priority should be given to expanding the range of costs considered in the planning process. For example, costs of recruitment strategies would be important in assessing screening participation while costs involved in training adequate numbers of people and providing adequate instruments to support colonoscopy follow-up of positive test results would be an important part of start-up costs.

References

Ackerman LV, Spjut HJ & Spratt JS 1964. The biological characteristics of colonic and rectal neoplasms with refutation of the concept that adenomatous polyps are highly premalignant tumours. *Acta Un Int Cancer* 20: 716-723.

Agresti A 1990. Categorical data analysis. New York: John Wiley & Sons.

Ahlquist DA, Wieand HS, Moertel CG, et al 1993. Accuracy of fecal occult blood screening for colorectal neoplasia. *JAMA* 269(10): 1262-1267.

Albert A, Gertman PM & Louis TA 1978. Screening for the early detection of cancer - I. The temporal natural history of a progressive disease state. *Mathematical Biosciences* 40: 1-59.

Albert A, Gertman PM, Louis TA & Liu S-I 1978. Screening for the early detection of cancer - II. The impact of the screening on the natural history of the disease. *Mathematical Biosciences* 40: 61-109.

Albert A 1981. Estimated cervical cancer disease state incidence and transition rates. *J Natl Cancer Inst* 67: 571-576.

Armitage P & Doll R 1961. Stochastic models for carcinogenesis. In: *Fourth Berkeley Symposium on Mathematical Statistics and Probability*. California: University of California Press, pp. 19-27.

Australian Health Ministers' Advisory Council (AHMAC) 1990. Breast cancer screening in Australia: future directions. Canberra: Australian Institute of Health: Prevention Program Evaluation Series No 1. AGPS.

Australian Health Ministers' Advisory Council (AHMAC) 1991. Cervical cancer screening in Australia: options for change. Canberra: Australian Institute of Health: Prevention Program Evaluation Series No 2. AGPS.

Australian Health Technology Advisory Committee 1997. Colorectal cancer screening. Canberra: Australian Government Publishing Service.

Australian Institute of Health and Welfare (AIHW) 1992. Australia's health 1992: the third biennial report of the Australian Institute of Health and Welfare. Canberra: AGPS.

Australian Institute of Health and Welfare (AIHW) 1997. Health expenditure bulletin number 13. Canberra: Australian Institute of Health and Welfare.

Australian Institute of Health and Welfare (AIHW) & Australasian Association of Cancer Registries (AACR) 1999. Cancer in Australia 1996: Incidence and mortality data for 1996 and selected data for 1997 and 1998. Canberra: AIHW (Cancer Series).

Baker SG, Connor RJ & Prorok PC 1991. Recent develpments in cancer screening modeling. In: Miller AB, Chamberlain J, Day NE, Hakama M & Prorok PC (eds.) *Cancer screening*. Cambridge: Cambridge University Press, pp. 404-418.

Bennett S, Stevenson C, Melville G, de Looper M & Wright P 1992. Mortality surveillance, Australia 1979-1990. Canberra: Australian Institute of Health and Welfare: Mortality Surveillance Series, No. 1, AGPS.

Bennett S, Donovan J, Stevenson C & Wright P 1994. Mortality surveillance, Australia 1981-1992. Canberra: Australian Institute of Health and Welfare: Mortality Surveillance Series, No. 2, AGPS.

Blumenson LE 1976. When is screening effective in reducing the death rate? *Mathematical Biosciences* 30: 273-303.

Blumenson LE 1977a. Detection of disease with periodic screening: Transient analysis and application to mammography examination. *Mathematical Biosciences* 33: 73-106.

Blumenson LE 1977b. Compromise screening strategies for chronic disease. *Mathematical Biosciences* 34: 79-94.

Brillinger DR 1986. The natural variability of vital rates and associated statistics. *Biometrics* 42: 693-734.

Brookmeyer R & Day NE 1987. Two-stage models for the analysis of cancer screening data. *Biometrics* 43(3): 657-669.

Bross IDJ, Blumenson LE, Slack NH & Priore RL 1968. A two disease model for breast cancer. In: Forrest APM & Bunkler PB (eds.) *Prognostic factors in Breast Cancer*. Baltimore: Williams and Wilkins, pp. 288-300.

Brown ML 1993. Screening for colorectal cancer (letter). *N Engl J Med* 329: 1352-1353.

Burt RW, Bishop DT, Lynch HT, Rozen P & Winawer SJ 1990. Risk and surveillance of individuals with heritable factors for colorectal cancer. *Bull World Health Organ* 68: 655-665.

Cali RL, Pitsch RM, Thorson AG, et al 1993. Cumulative incidence of metachronous colorectal cancer. *Dis Colon Rectum* 36: 388-393.

Chiang CL 1964. A stochastic model of competing risks of illness and competing risks of death. In: Gurland J (ed.) *Stochastic models in medicine and biology*. Madison: University of Wisconsin Press, pp. 323-354.

Chiang CL 1968. Introduction to stochastic processes in biostatistics. New York London Sydney: John Wiley & Sons, Inc.

Chiang CL 1980. An introduction to stochastic processes and their applications. Huntington, New York: Robert E. Krieger Publishing Company.

Chu KC & Connor RJ 1991. Analysis of the temporal patterns of benefits in the Health Insurance Plan of Greater New York Trial by stage and age. *Am J Epidemiol* 133(10): 1039-1049.

Clark ML, Price AB & Williams CB 1987. Tumours of the gastrointestinal tract. In: Weatherall DJ, Ledingham JGG & Warrell DA (eds.) *Oxford textbook of medicine*. Oxford, Melbourne, New York: Oxford University Press, pp. 12.156-12.157.

Collette HJA, Day NE, Rombach JJ & de Waard F 1984. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case control study. *Lancet* I: 1224-1226.

Colombo L, Corti G, Magri F, et al 1997. Results of a pilot study of endoscopic screening of first degree relatives of colorectal cancer patients in Italy. *J Epidemiol Community Health* 51(4): 453-458.

Connor RJ, Chu KC & Smart CR 1989. Stage-shift cancer screening model. *J Clin Epidemiol* 42: 1083-1095.

Cooper GS, Yuan Z, Landefeld CS, Johanson JF & Rimm AA 1995. A national population-based study of incidence of colorectal cancer and age. *Cancer* 75: 775-781.

Coppleson LW & Brown B 1975. Observations on a model of the biology of carcinoma of the cervix: a poor fit between observations and theory. *American Journal of Obstetrics and Gynecology* 122: 127-136.

Day NE & Walter SD 1984. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 40(1): 1-14.

Day NE & Miller AB 1988. Screening for breast cancer. Toronto: Hans Huber Publishers.

Day NE 1990. The Armitage-Doll multistage model of carcinogenesis. *Stat Med* 9: 677-679.

Day NE & Duffy SW 1996. Trial design based on surrogate end points—application to comparison of different breast screening frequencies. *J R Statist Soc A* 159 (1): 49-60.

Du Pasquier 1913. Mathematische theorie der invaliditatsversicherung. *Milt Verein Schweiz Versich -Math* 8: 1-153.

Dubin N 1979. Benefits of screening for breast cancer: Application of a probabilistic model to a breast cancer detection project. *J Chronic Dis* 32: 145-151.

Dubin N 1981. Predicting the benefit of screening for disease. *J Appl Prob* 18: 348-360.

Drummond MF, Stoddart GL & Torrance GW 1987. Methods for the economic evaluation of health care programmes. Oxford, New York, Toronto: Oxford University Press, p. 43.

Eddy DM 1980. Screening for cancer: theory, analysis and design. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Eddy DM 1984. The cost-effectiveness of colorectal cancer screening. In: Levin B & Riddell R (eds.) *Gastrointestinal cancer*. New York: Elsevier Science Publishing,

Eddy DM 1985. Technology assessment: the role of mathematical modeling. In: Institute of Medicine (ed.) *Assessing medical technologies*. Washington, D.C.: National Academy Press, pp. 144-153.

Eddy DM 1987. The frequency of cervical cancer screening. Comparison of a mathematical model with empirical data. *Cancer* 60(5): 1117-1122.

Eddy DM, Nugent FW, Eddy JF, et al 1987. Screening for colorectal cancer in a high-risk population. *Gastroenterology* 92: 682-692.

Eddy DM 1990. Screening for colorectal cancer. *Ann Intern Med* 113(5): 373-384.

Fenoglio CM & Lane N 1974. The anatomical precursor of colorectal carcinoma. *Cancer* 34: 819-823.

Figiel LS, Figiel SJ & Wietersen FK 1965. Roentgenologic observations of growth rates of colonic polyps and carcinoma. *Acta Radiologica Diagnosis* 3: 417-428.

Fix E & Neyman J 1951. A simple stochastic model of recovery, relapse, death and loss of patients. *Human Biol* 23: 205-241.

Fleischer DE, Goldberg SB, Browning TH, et al 1989. Detection and surveillance of colorectal cancer. *JAMA* 261: 580-585.

Frommer DJ, Kapparis A & Brown MK 1988. Improved screening for colorectal cancer by immunological detection of occult blood. *BMJ* 296: 1092-1094.

Fuchs CS, Giovannucci EL, Colditz GA, Hynter D/\J, Speizer FE & Willett WC 1994. A prospective study of family history and the risk of colorectal cancer. *N Engl J Med* 331(25): 1669-1674.

George B, Harris A & Mitchell A 1998. Reimbursement decisions and the implied value of life: Cost effectiveness analysis and decisions to reimburse pharmaceuticals in Australia 1993-1996. In: Harris AH (ed.) *Economics and Health: 1997. Proceedings of the 19th Australian Conference of Health Economists. Australian Studies in Health Science*

Administration No. 85. Sydney: School of Health Sciences Management, University of New South Wales, pp. 1–17.

Graham P 1996. Sigmoidoscopy. *Australian Family Physician* 25: 1403-1404.

Greegor DH 1967. Diagnosis of large-bowel cancer in the asymptomatic patient. *JAMA* 201: 943-945.

Gyrd-Hansen D, Sogaard J & Kronborg O 1997. Analysis of screening data: colorectal cancer. *Int J Epidemiol* 26: 1172-1181.

Habbema JDF, Lubbe JThN, van der Maas PJ & van Oortmarssen GJ 1983. A computer simulation approach to the evaluation of mass screening. In: van Bemmel et al (ed.) *MEDINFO 83. Proceedings of the 4th World Conference on Medical Informatics*. Amsterdam: North Holland, pp. 1222.

Harbison SP & Steele GL 1991. C, A reference manual. 3rd Ed. Englewood Cliffs, New Jersey: Prentice Hall.

Hardcastle JD, Farrands PA, Balfour TW, Chamberlain J, Amar SS & Sheldon MG 1983. Controlled trial of faecal occult blood testing in the detection of colorectal cancer. *Lancet* ii: 1-4.

Hardcastle JD, Chamberlain J, Sheffield J, et al 1989. Randomised controlled trial of faecal occult blood screening for colorectal cancer: Results for first 107,349 subjects. *Lancet* i: 1160-1164.

Hardcastle JD, Chamberlain JO, Robinson MHE, et al 1996. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 348: 1472-1477.

Hoff G, Foerster A, Vatn MH, Sauar J & Larsen S 1986. Epidemiology of polyps in the rectum and colon. *Scand J Gastroenterol* 21: 853-862.

Houlston RS, Murday V, Harocopos C, Williams CB & Slack J 1990. Screening and genetic counselling for relatives of patients with colorectal cancer in a family clinic. *BMJ* 301(6748): 366-368.

Jelfs P, Coates M, Giles G, et al 1996. Cancer in Australia 1989-1990 (with projections to 1995). Canberra: Australian Institute of Health and Welfare (Cancer Series No. 5).

Kernighan BW & Ritchie DM 1988. The C programming language. 2nd Ed. Englewood Cliffs, New Jersey: Prentice Hall Software Series.

Kewenter J, Bjork S, Haglind E, Smith L, Svanvik J & Ahren C 1988. Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27,700 subjects. *Cancer* 62: 645-651.

Kewenter J, Asztely M, Engaras B, Haglind E, Svanvik J & Ahren C 1991. A randomised trial of fecal occult blood testing for early detection of colorectal cancer. Results of screening and rescreening 51325 subjects. In: Miller AB, Chamberlain J, Day NE, Hakama M & Prorok PC (eds.) *Cancer screening*. Cambridge: Cambridge University Press, pp. 116-125.

Knight KK, Fielding JE & Battista RN 1989. Occult blood screening for colorectal cancer. *JAMA* 261(4): 587-593.

Knoernschild HE 1963. Growth rate and malignant potential of colonic polyps: early results. *Surgical Forum* 14: 137-138.

Knox EG 1973. A simulation system for screening procedures. In: McLachlan G (ed.) *Future and present indicatives, problems and progress in medical care, ninth series*. Oxford, London: Nuffield Provincial Hospitals Trust, pp. 17-55.

Knox EG 1975. Simulation studies of breast cancer screening programmes. In: McLachlan G (ed.) *Probes for health*. London, New York, Toronto: Oxford University Press, pp. 13-44.

Knox EG 1988. Evaluation of a proposed breast cancer screening regimen. *BMJ* 297: 650-654.

Knox EG & Woodman CBJ 1988. Effectiveness of a cancer control programme. *Cancer Surveys* 7 (3): 379-401.

Kozuka S, Nogaki M, Ozeki T & Masumori S 1975. Premalignancy of the mucosal polyp in the large intestine. II. estimation of the periods required for malignant transformation of mucosal polyps. *Dis Colon Rectum* 18: 494-500.

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.

Lang CA & Ransohoff DF 1994. Fecal occult blood screening for colorectal cancer—is mortality reduced by chance selection for screening colonoscopy. *JAMA* 271: 1011-1013.

Lieberman D 1995. Cost-effectiveness model for colon cancer screening. *Gastroenterology* 109: 1781-1790.

Lincoln T & Weiss GH 1964. A statistical evaluation of recurrent medical examination. *Operations Research* 12: 187-205.

Loeve F, Brown ML, Boer R, van Ballegooijen M, van Oortmarssen GJ, Habbema JDF 2000. Endoscopic colorectal cancer screening: a cost-saving analysis. *J Natl Cancer Inst* 92: 557–63.

Louis TA, Albert A & Heghinian S 1978. Screening for the early detection of cancer - III. Estimation of disease natural history. *Mathematical Biosciences* 40: 111-144.

Macrae FA 1996. Screening for colorectal cancer, 1996. *Med J Aust* 165: 102-105.

Mandel JS, Bond JH, Church TR, et al 1993. Reducing mortality from colorectal cancer by screening for fecal occult blood. *N Engl J Med* 328: 1365-1371.

Mandel JS, Ederer F, Church T, et al 1994. Screening for colorectal cancer: which test is best? *JAMA* 272:1099.

Marcus AC, Ahnen D, Cutter G, et al 1999. Promoting cancer screening among first-degree relatives of breast and colorectal cancer patients: the design of two randomised trials. *Prev Med* 28: 229-242.

Martyres R, St John DJB, Irving FH & Wyman K 1999. Colorectal cancer screening in general practice: A survey of current practice and attitudes in Victoria. *Australian Family Physician* 28(7): 755-758.

Mathsoft Inc 1995. S-plus User's Manual version 3.3. Seattle, Washington: Mathsoft, Inc.

McLeod AI 1985. A remark on Algorithm AS 183. An efficient and portable pseudo-random number generator. *Appl Statist* 34: 198-200.

Miller AB, Chamberlain J, Day NE, Hakema M & Prorok PC 1990. Report on a workshop of the UICC Project on evaluation of screening for cancer. *Int J Cancer* 46(5): 761-769.

Morrison AS 1985. Screening in chronic disease. New York Oxford: Oxford University Press.

Morson BC 1974. Evolution of cancer of the colon and rectum. *Cancer* 34: 845-849.

Muto T, Bussey HJR & Morson BC 1975. The evolution of cancer of the colon and rectum. *Cancer* 36: 2251-2270.

National Health and Medical Research Council (NHMRC) 1999. Guidelines for the prevention, early detection and management of Colorectal Cancer (CRC). Canberra: NHMRC.

Newland RC, Chapuis PH, Pheils MT & MacPherson JG 1981. The relationship of survival to staging and grading of colorectal carcinomas. *Cancer* 47: 1424-1429.

Nicholson FB, Korman MG, Stern AI & Hansky J 2000. Distribution of colorectal adenomas: implications for bowel cancer screening. *Med J Aust* 172: 428–430.

Niv Y 1992. Does a risk questionnaire add anything to a colorectal screening project? Report of a three year screening experience. *J Clin Gastroenterol* 15(1): 33-36.

Norfleet RG 1986. Effect of diet on fecal occult blood testing in patients with colorectal polyps. *Dig Dis Sci* 31: 498-501.

O'Brien MJ, Winawer SJ, Zauber AG, et al 1990. The National Polyp Study: Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. *Gastroenterology* 98: 371-379.

O'Neill TJ, Tallis GM & Leppard P 1995. A review of the technical features of breast cancer screening illustrated by a specific model using South Australian cancer registry data. *Statistical Methods in Medical Research* 4: 55-72.

Office of the Australian Government Actuary 1995. Australian Life Tables 1990-1992. Canberra: Australian Government Publishing Service.

Olynyk JK, Aquilia S, Fletcher DR & Dickinson JA 1996. Flexible sigmoidoscopy screening for colorectal cancer in average-risk subjects: a community-based pilot project. *Med J Aust* 165: 74-76.

Parkin DM 1985. A computer simulation model for the practical planning of cervical cancer screening programmes. *Br J Cancer* 51(4): 551-568.

Prorok PC 1976a. The theory of periodic screening II: Doubly bounded recurrence times and mean lead time and detection probability estimation. *Advances in Applied Probability* 8: 460-476.

Prorok PC 1976b. The theory of periodic screening I: Lead time and proportion detected. *Advances in Applied Probability* 8: 127-143.

Prorok PC 1986. Mathematical models and natural history in cervical cancer screening. In: Hakama M, Miller AB & Day NE (eds.) *Screening for cancer of the uterine cervix.* IARC Scientific Publication 76, pp. 185-198.

Prorok PC 1988. Mathematical models of breast cancer screening. In: Day NE & Miller AB (eds.) *Screening for breast cancer.* Toronto: Hans Huber Publishers, pp. 95-109.

Rae LC 1998. Community screening for colorectal cancer in north-eastern New South Wales, 1987-1996. *Med J Aust* 168: 382-385.

Rex DK, Lehman GA, Hawes RH, Ulbright TM & Smith JJ 1991. Screening colonoscopy in asymptomatic average-risk persons with negative fecal occult blood tests. *Gastroenterology* 100: 64-67.

Ripley BD 1987. Stochastic simulation. New York: John Wiley & Sons.

Robinson MHE, Moss SM, Whynes DK, Chamberlain JO, Hardcastle JD & Mangham CM 1995. The effect of dietary restriction retesting on Haemoccult sensitivity during population screening for colorectal carcinoma. *J Med Screening* 2: 41-44.

Salkeld G, Young G, Irwig L, Haas M & Glasziou P 1996. Cost-effectiveness analysis of screening by faecal occult blood testing for colorectal cancer in Australia. *Aust N Z J Public Health* 20: 138-143.

Schrage L 1979. A more portable FORTRAN random number generator. *ACM Transactions on Mathematical Software* 5: 132-138.

Selby JV & Friedman GD 1989. Sigmoidoscopy in the periodic health examination of asymptomatic adults. *JAMA* 261(4): 595-601.

Selby JV, Friedman GD, Quesenberry CP & Weiss NS 1992. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med* 326: 653-657.

Shapiro S, Venet W, Strax P, Venet L & Roeser R 1982. Ten to fourteen year affect of screening on breast cancer mortality. *J Natl Cancer Inst* 69: 349-355.

Shimbo T, Glick HA & Eisenberg JM 1994. Cost-effectiveness analysis of strategies for colorectal cancer screening in Japan. *Int J Technol Assess Health Care* 10: 359-375.

Shwartz M 1978a. An analysis of the benefits of serial screening for breast cancer based upon a mathematical model of the disease. *Cancer* 41: 1550-1564.

Shwartz M 1978b. A mathematical model used to analyse breast cancer screening strategies. *Operations Research* 26 (6): 937-955.

Shwartz M & Plough A 1984. Models to aid in planning cancer screening programs. In: Cornell RG (ed.) Statistical methods for cancer studies. New York and Basel: Marcel Dekker, Inc., pp. 239-416.

South Australian Health Commission & South Australian Cancer Registry 1994. Epidemiology of cancer in South Australia 1977 to 1993. Adelaide: South Australian Cancer Registry.

St. John DJB, McDermott FT, Hopper JL, Debney EA, Johnson WR & Hughes ESR 1993. Cancer risk in relatives of patients with common colorectal cancer. *Ann Intern Med* 118: 785-790.

St. John DJB, Young GP, Alexeyeff MA, et al 1993. Evaluation of new occult blood tests for detection of colorectal neoplasia. *Gastroenterology* 104:1661-1668.

Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM & MacCarty RL 1987. Natural history of untreated colonic polyps. *Gastroenterology* 93: 1009-1013.

Tabar L, Fagerberg G, Duffy S & Day NE 1989. The Swedish two county trial of mammography screening for breast cancer: recent results and calculation of benefit. *Journal of Community Health* 43: 107-114.

Tallis GM, Leppard P & O'Neill TJ 1988. The analysis of survival data from a central cancer registry with passive follow-up. *Stat Med* 7: 483-490.

Tazi MA, Faivre J, Dassonville F, Lamour J, Milan C & Durand G 1997. Participation in faecal occult blood screening for colorectal cancer in a well defined French population: results of five screening rounds from 1988 to 1996. *J Med Screening* 4: 147-151.

Thomas D 1992. Use of Biologic Models to Analyze Epidemiologic Data, Application: Stochastic Models of Carcinogenesis, Gene-Environment Interactions, and Breast and Lung Cancer. In: Dwyer JH, Feinleib M, Lippert P & Hoffmeister H (eds.) *Statistical Models for Longitudinal Studies of Health*. New York Oxford: Oxford University Press, pp. 351-366.

Thomas W, White CM, Mah J, Geisser MS, Church TR & Mandel JS 1995. Longitudinal compliance with annual screening for fecal occult blood. *Am J Epidemiol* 142(2): 176-182.

Thompson JR & Brown BW 1987. Cancer modeling. New York and Basel: Marcel Dekker, Inc.

van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al 1990. A model for breast cancer screening. *Cancer* 66(7): 1601-1612.

van Oortmarssen GJ & Habbema JD 1991. Epidemiological evidence for age-dependent regression of pre-invasive cervical cancer. *Br J Cancer* 64(3): 559-565.

Verbeek ALM, Straatman H & Hendriks JHCL 1988. Sensitivity of mammography in Nijmegen women under age 50: some trials with the Eddy model. In: Day NE & Miller AB (eds.) *Screening for Breast Cancer*. Toronto: Hans Huber Publishers, pp. 29-38.

Wagner JL, Herdman RC & Wadhwa S 1991. Cost effectiveness of colorectal cancer screening in the elderly. *Ann Intern Med* 115(10): 807-817.

Wagner JL, Tunis S, Brown M, Ching A & Almeida R 1996. Cost-effectiveness of colorectal cancer screening in average-risk adults. In: Young G & Levin B (eds.) *Prevention and early detection of colorectal cancer*. London: Saunders, pp. 321-356.

Walter SD & Day NE 1983. Estimation of the duration of a preclinical state using screening data. *Am J Epidemiol* 118: 865-886.

Walter SD & Stitt LW 1987. Evaluating the survival of cancer cases detected by screening. *Stat Med* 6: 885-900.

Weinstein MC, Siegel JE, Gold MR, Kamlet MS & Russell LB 1996. Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA* 276: 1253-1258.

Welin S, Youker J & Spratt JS 1963. The rates and patterns of growth of 375 tumors of the large intestine and rectum observed serially by double contrast enema study (Malmo technique). *Am J Roentgenol, Rad Therapy & Nuclear Med* 90(4): 673-687.

Wichmann BA & Hill ID 1982. Algorithm AS 183. An efficient and portable pseudo-random number generator. *Appl Statist* 31: 188-190.

Wichmann BA & Hill ID 1983. An efficient and portable pseudo-random number generator (correction). *Appl Statist* 32: 123.

Winawer SJ, O'Brien MJ, Waye JD, et al 1990. Risk and surveillance of individuals with colorectal polyps. *Bull World Health Organ* 68 (6): 789-795.

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

Zeisel H 1986. A remark on Algorithm AS 183. An efficient and portable pseudo-random number generator. *Appl Statist* 35: 89.

Zelen M & Feinleib M 1969. On the theory of screening for chronic diseases. *Biometrika* 56: 601-614.

Appendix A: The computer program

The disease and screening model was implemented in a computer program written in ANSI standard C mainly on a UNIX computer. The final program consisted of approximately 8,000 lines of code. This appendix describes key features of the program.

Implementation language

At the time this project was started there was no readily available, existing software for microsimulation modelling. Hence any such software needed to be written before the modelling could be applied. Given the time constraints of a PhD, I needed a language that was easy to learn and simple to apply. The main criteria for the choice of implementation language were:

- The package or language should be efficient

A microsimulation model involves relatively simple calculations applied to a large number of records. Hence a high level language or package such as SAS which contains sophisticated analysis routines and simplified data handling but which requires extra time and computer resources to store and manipulate data is less appropriate than a lower level language such as FORTRAN or C.

- The package or language should be widely available

The model was developed and used on a number of different computers, including larger UNIX computers and smaller PC's. Further, if the final model is to be used by policy makers and planners, it should be written in a widely available language. The growing number of C language programs and programmers along with the wide acceptance of a standard form of this language means that it easily meets this criterion. (Harbison & Steele 1991)

- The package or language should be suitable for both small and large data sets.

Microsimulations can be used for a variety of populations ranging from relatively small, for example modelling screening for a specific age cohort, to very large, for example modelling a national cancer screening program. Hence the model needs to handle small and large data sets equally well. An advantage of higher level packages such as SAS is that their more sophisticated data handling capabilities allow for efficient storage of data sets of varying sizes. Older programming languages such as FORTRAN require memory to be set aside for the largest data set the program will meet, leading to inefficient computer memory use. However, modern languages such as C or PL1 allow for dynamic memory allocation so that

computer memory requirements are dictated by the data storage required in any particular case.

- The structure of the language or package should be simple

The C language has relatively few commands, can be described in a small space and learned quickly. (Kernighan & Ritchie 1988) Hence it was possible to learn it in a relatively short time to a standard which allowed the writing of fairly complex and sophisticated code.

I chose to use the C language because it was readily available and met these criteria (Harbison & Steele 1991). Hence the first part of my PhD course was spent learning how to use the C language and developing the software tools for implementing the microsimulation model.

Program structure

The program is written in a modular style. Each specific task is written in a separate module of code, with the flow of tasks being governed by a relatively simple main program. The reason for this is that the final model evolved over the duration of the PhD project and is likely to evolve further as new data become available for model parameter specification and estimation. The modular style of program ensures that changes in any part of the model should be confined to a small number of program modules, allowing relatively easy modification. Figure A1 shows the basic program structure.

The program starts by reading the preclinical incidence rates, adjusted death probabilities and population counts. The preclinical incidence and adjusted death rates are calculated separately, in a Microsoft Exel spreadsheet, using the formulae specified in chapter 3. These rates and counts are classified by age in single years and by sex. The program also reads a file of screening participation rates and a file of control parameters which specify the screening protocol—for example, the age ranges to be screened, the period over which the screening program will be simulated, etc.

The program then sets up the simulated population as a linked chain of C structures, where each node on the chain represents a population member. Memory is assigned to each member as it is added to the chain, allowing the program to assign only such memory as is needed for the specific simulated population. Figure A2 is a listing of the C code which generates each population member.

This is followed by the program modules which randomly allocate cancers to each member according to the pre-clinical incidence rates. Polyps are added as a linked chain to the pointer `p->polyp`. This chain specifies whether or not the polyp will progress to a cancer and, if so, the age at which the cancer will surface, whether or not it will be cured and, if not, the survival time from surfacing to death. Figure A3 is a listing of the C code which adds a polyp to the polyp chain and figure A4 is a listing of the code which applies the cancer model to the polyp.

The program allows the user to specify a number of iterations for the whole simulation. For each iteration, a synthetic population is generated, the screening simulation applied, the results written to the output structure and the population memory freed before the next iteration. In principle the whole synthetic population could be simulated at once. However, each computer imposes limits on the total amount of memory available to any specific user. Hence, the population had to be generated separately for each age cohort, the screening simulation applied and then the memory freed before moving on to the next age cohort in order for the program to be run on the University Unix computer.

When the program has finished all iterations of the simulation, the results are written to an overall summary output file and two individual year results files.

Screening module

The main program calls the screening module to apply the screening model to the synthetic population. Figure A5 shows the structure of this screening module.

The screening module is applied to each population member in turn. First the control group results are added to the aggregate results structure. These are the numbers of cancers and the associated deaths, cures and costs in the absence of screening. Then the member is checked to see if the age lies above the screening program age range or if an existing cancer has clinically surfaced. If so, the member is ineligible for screening, so the results are added to the screening group aggregate results structure and the module moves to the next population member.

If the population member is eligible for screening, then the age is incremented to the lowest screening age range (if necessary) and the screening model is applied. If a cancer or polyp is found, then the population unit leaves the screening program and the polyp/cancer treatment and follow-up model is applied. If no cancer or polyp is found and if the population unit survives death from other causes, then ultimately the age is incremented to the highest screening age range. At this point any further cancers arising in the screening period are added to the aggregate screening group results structure and the module moves on to the next population member.

At each point in this process where the member age is incremented, the time elapsed from the start of the simulation is checked to see if it is still within the screening period. Once the end of the screening period is reached, no further results are added to either the control or screening group results structure, so that only cancers arising or deaths occurring in the screening period are included.

Figures A6 and A7 list the C code for the modules which apply FOBT screening and determine the outcome. The modules applying sigmoidoscopy and colonoscopy screening follow a similar pattern. The first module checks the polyp chain attached to each population member to look for any polyps or cancers which exist at the time of screening. The subroutines `screen_pol()` and `screen_canc()` then check if the polyp or cancer is found by the screen according to the sensitivity implied by the polyp size or the cancer stage. The results are then recorded as either a true positive,

true negative, false positive or false negative. The parameter `SPECFOBT` is the probability of a false positive. The parameter `DECLINE` allows a proportion of people to decline follow up to screening, in which case a true positive becomes a false negative and a false positive becomes a true negative. The function `genscrn()` returns a positive random number less than one. The random number generator is discussed in detail in Appendix B.

The second module code reads the screen result (true or false, positive or negative). If the result was a true positive, it sets flags to show that a colonoscopy and a clinical examination were also performed. These will be used in calculating the cost of the screen. It also checks to see if the colonoscopy resulted in a perforation. The module reads the polyp chain to find the largest polyp or most advanced cancer present at the time of screening. This is deemed to be the stage of the screen detected cancer or polyp. The function `detect_cancer()` determines the outcome (in terms of cure rate and survival time) for the screen detected cancer.

The stage at detection is forced to be no greater than the stage at clinical surfacing in the absence of screening and if the stage at detection is the same as the stage at clinical surfacing, then the prognosis is forced to be the same. Usually the detection of a polyp would lead to a person leaving the screening program and being monitored with regular colonoscopies for further polyps. However, the parameter `SMPOLYP` allows me to specify a polyp size below which the polyp is removed but the person goes back into the regular screening program.

If the result is a false positive, the module records a colonoscopy, a clinical examination and, where appropriate, a perforation for use in the cost estimate. Where the results are either true or false negatives, the program simply records that the person continues in the screening program.

Figures

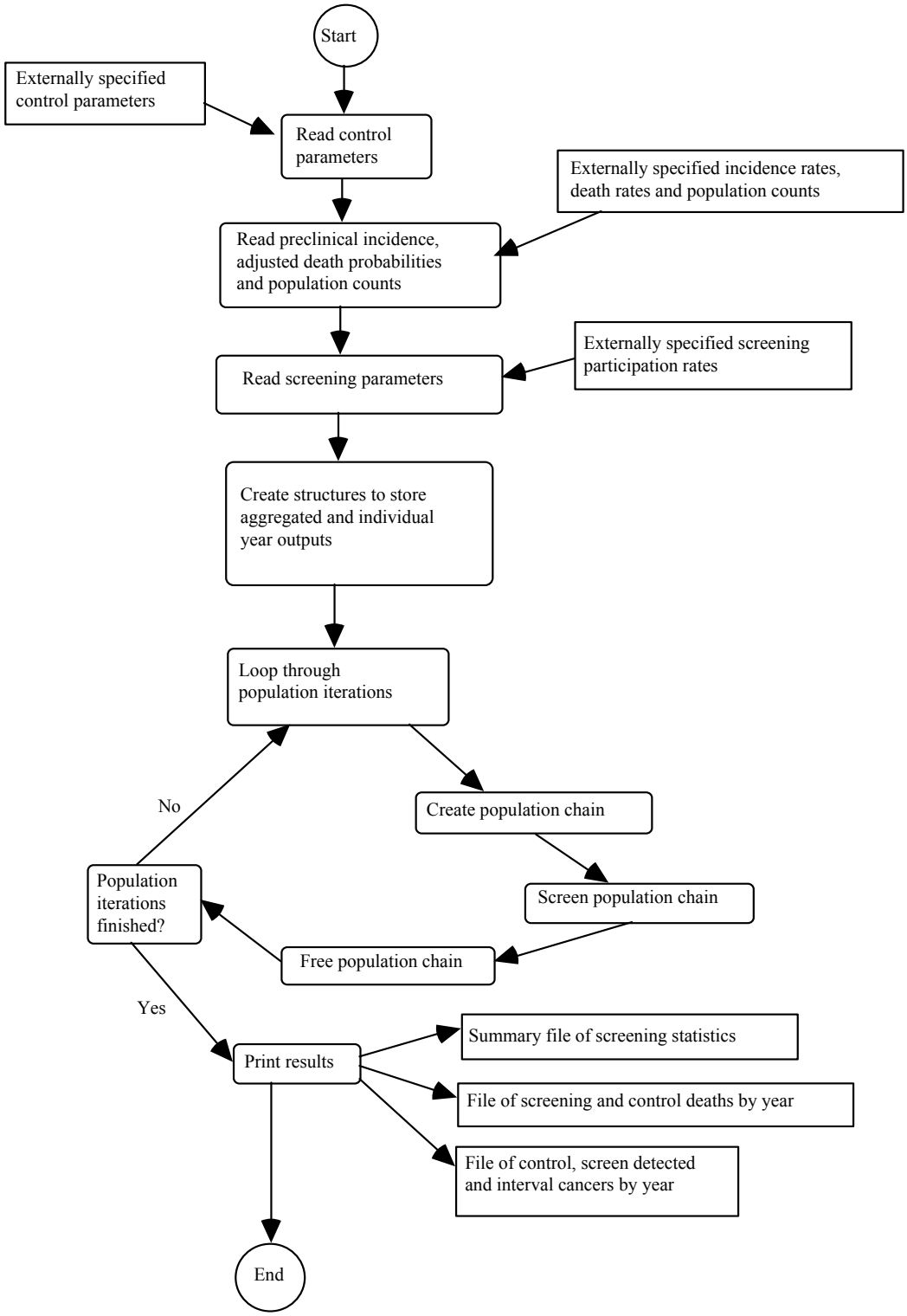

Figure A1: Schematic diagram of main program

```
/*****************************************************************

   addmem

   This function adds a member to the population chain and returns a
   pointer to this member. It calls the function malloc to allocate
   the space for the member.

   Function inputs:
    ptr - pointer to the last member of the chain
    age - current age to add members

   The function returns a pointer to the new member.

*****************************************************************/
struct member *addmem(struct member *plast,
                      int age)
{
    struct member *p;
     p = (struct member *) malloc(sizeof(struct member));
     if (p==NULL) {
        printf("no space for member at age %d\n", age);
        exit(1);
     }
     p->age_start = (double) age + genpopn();
     p->age_death = MAXAGE;
     p->age_surface = MAXAGE;
     p->stage_surface = NOTDET;
     p->canc_death = MAXAGE;
     p->age_detect = MAXAGE;
     p->detect_death = MAXAGE;
     p->stage_detect = NOTDET;
     p->perf_death = NO;
     p->canc_cure = YES;
     p->screen_cure = YES;
     p->back_scr = NO;
     p->back_cost = 0.0;
     p->back_cost_disc = 0.0;
```

Figure A2: C routine to allocate member of the population chain

```
      p->new_polyp = NO;
      p->risk = AVERAGE;
      p->ppolyp = NULL;
      p->age_now = 0;
      p->time_now = 0;
      p->pscreen = NULL;
      p->last = plast;
      p->next = NULL;
      return p;
}
```

Figure A2: C routine to allocate member of the population chain *(continued)*

Notes:

1. MAXAGE is the oldest age included in the simulation (100 years)
2. Setting the risk to AVERAGE, identifies the member as not in a risk group. Allocation to a risk group is done elsewhere in the program.
3. Setting the stage to NOTDET indicates that whether or not the unit has cancer has not been determined. This is determined elsewhere in the program.

```
/*****************************************************************

     addpol

     This function adds a member to the polyp chain and runs the
     cancer model on it.

*****************************************************************/
struct polyp *addpol(struct polyp *plast,
                double age,
                int sex,
                double incidence[][MAXAGE+1][4],
                int outcome)
{
     double prob;
     double polsize=POLSIZE;
     int iage=floor(age);
     int i;
     struct polyp *p;

     p = (struct polyp *) malloc(sizeof(struct polyp));

     if (p==NULL) {
      printf("no space for polyp at age %d\n", age);
      exit(1);
     }
     p->size = lognormal(POLMEAN, POLVAR);
     p->outcome = outcome;
     prob = genpopn();
     if (outcome == CERTAIN) {
         if (prob <= incidence[sex][iage][PROXIMAL])
          p->site = PROXIMAL;
         else if (prob <= (incidence[sex][iage][DISTAL]
                    + incidence[sex][iage][PROXIMAL]))
          p->site = DISTAL;
         else
          p->site = RECTAL;
     }
```

Figure A3: C code to add a polyp to the polyp chain

```
     else {
       if (prob <= PROXPT)
          p->site = PROXIMAL;
       else if (prob <= DISTPT)
          p->site = DISTAL;
         else
          p->site = RECTAL;
     }

     p->age_onset = age;
     p->age_dwell = exponential(1/PDWELL) + age;
     if (p->size < polsize)
        p->age_pol10 = p->age_dwell;
     else
      p->age_pol10 = age;
     for (i=0; i<4; i++)
         p->age_stage[i] = MAXAGE;
     if (p->outcome == 1)
      p->age_surface = cancer(p, sex);
     else {
        p->growth = 0;
      p->size_transit = 0;
      p->age_surface = MAXAGE;
        p->age_death = MAXAGE;
        p->stage_surface = NOTDET;
        p->cure = YES;
     } /* outcome not cancer */
     p->last = plast;
     p->next = NULL;

     return p;
}
```

Figure A3: C code to add a polyp to the polyp chain ***(continued)***

```
/*****************************************************************

     cancer

     This function runs the cancer model for a specified polyp

     Function inputs:
      ptr - pointer to the current polyp
      (all other inputs are on the member pointed to by this
pointer)

     The function returns the age of surfacing.

*****************************************************************/
double cancer(struct polyp *ptr,
            int sex)
{
     int i;
     double age_surface;
     double prob;
     double polsize = POLSIZE;

     ptr->growth = beta(agrowth[sex][ptr->site],
                  bgrowth[sex][ptr->site], UPPERGROWTH,
LOWERGROWTH);
     ptr->size_transit = weibull_cond(LAMDATRANSIT, GAMMATRANSIT,
                     UPPERTRANSIT, LOWERTRANSIT);
     ptr->age_stage[0] = ptr->age_dwell +
           log(ptr->size_transit/ptr->size)/(ptr->growth*365.25);
     if(ptr->size < polsize && ptr->size_transit >= polsize)
      ptr->age_pol10 = ptr->age_dwell +
           log(polsize/ptr->size)/(ptr->growth*365.25);
     else if (ptr->size_transit < polsize)
      ptr->age_pol10 = ptr->age_stage[0];
     else
      ptr->age_pol10 = ptr->age_onset;

     prob = genpopn();
     if (prob <= surface[sex][ptr->site][0])
      ptr->stage_surface = 0;
```

Figure A4: C code to apply the cancer model to the polyp

```
     else if (prob <= surface[sex][ptr->site][1])
      ptr->stage_surface = 1;
     else if (prob <= surface[sex][ptr->site][2])
      ptr->stage_surface = 2;
     else
      ptr->stage_surface = 3;

     for (i=0; i<ptr->stage_surface; i++) {
       ptr->age_stage[i+1] = ptr->age_stage[i] +
exponential(sojourn[i]);
     }
     age_surface = ptr->age_stage[ptr->stage_surface] +
                exponential(sojourn[ptr->stage_surface]);
     for (i=0; i<=ptr->stage_surface; i++) {
      if (ptr->age_stage[i] >= MAXAGE)
         ptr->age_stage[i] = MAXAGE;
     }
     if (age_surface >= MAXAGE) {
       age_surface = MAXAGE;
       ptr->age_death = MAXAGE;
       ptr->cure = YES;
     }
     else {
         ptr->age_death = survival(ptr->stage_surface, CONTRL,
            &(ptr->cure), age_surface);
        if ( ptr->age_death >= MAXAGE)
           ptr->age_death = MAXAGE;
     }

     return age_surface;
}
```

Figure A4: C code to apply the cancer model to the polyp ***(continued)***

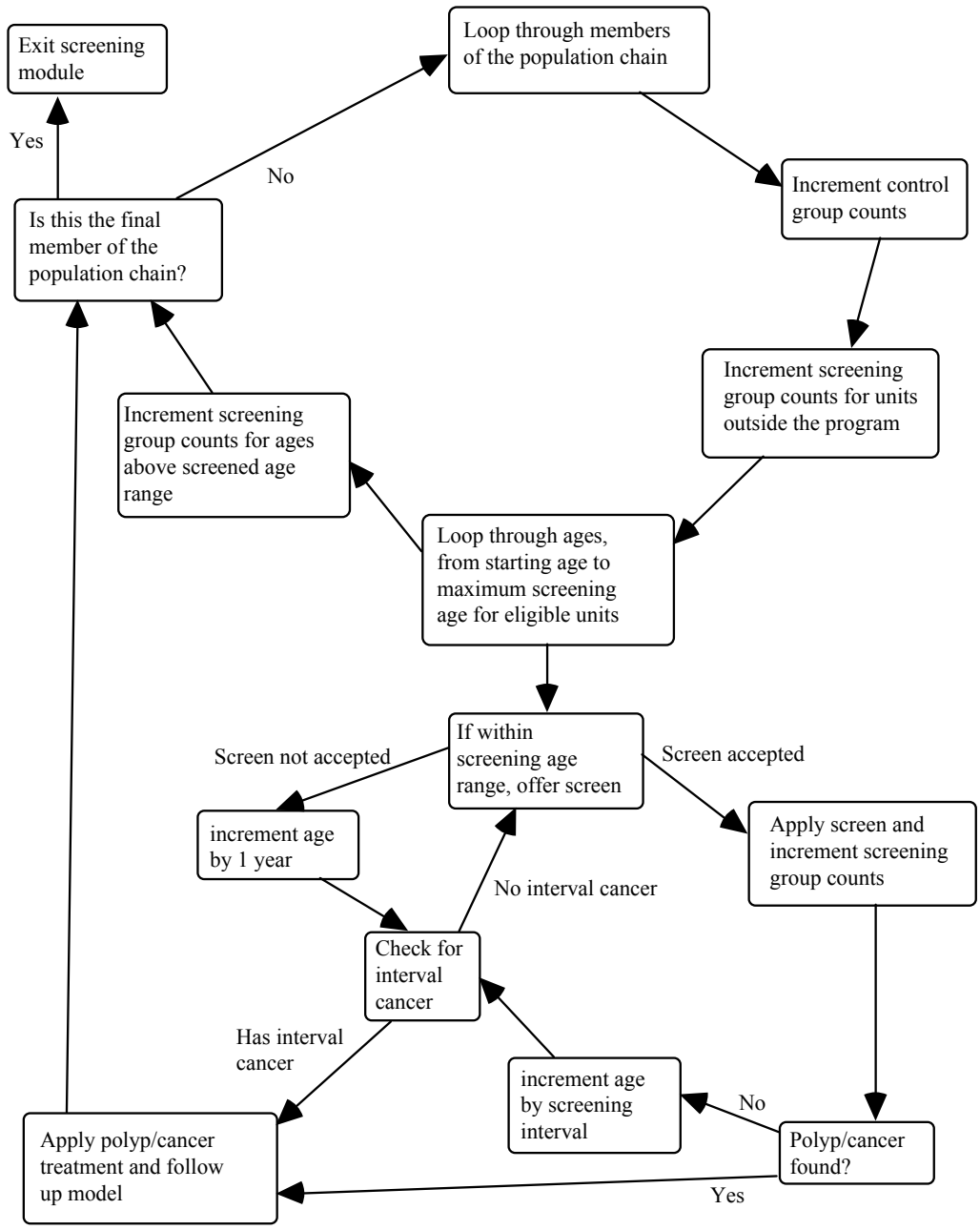

Figure A5: Schematic diagram of screening module

```
/*****************************************************************

    fobt
    Function performs the fobt screen and records the outcome.

    inputs:
    ptr - pointer to head of the polyp chain
    age - age at screening
    calls routines:
    screen_pol
    polsize
    screen_canc
    returns
    int result - screening result

*****************************************************************/
int fobt(struct member *ptr,
       double age)
{
     struct polyp *pol, *qpol;
     int result = NORES;
     int any = NO;
     int found = NO;
     double size;
     double prob;
     for(pol=ptr->ppolyp; pol!=NULL && found==NO; pol=qpol) {
      if(pol->age_onset <= age) {
         if(pol->outcome == 0) {
            if(pol->age_dwell >= age) {
               any = YES;
               found = screen_pol(pol->size);
               qpol = pol->next;
            }
            else
             qpol = pol->next;
         } /* polyp does not progress to cancer */
         else if(pol->age_dwell > age) {
              any = YES;
            found = screen_pol(pol->size);
```

Figure A6: C code which applies FOBT screen

```
            qpol = pol->next;
         } /* still polyp at initial size */
         else if(pol->age_stage[A] > age) {
              any = YES;
            size = polsize(pol,age);
            found = screen_pol(size);
            qpol = pol->next;
         } /* still polyp but growing */
         else {
            any = YES;
            found = screen_canc(pol, age);
            qpol = pol->next;
         } /* made transition to cancer */
        } /* if polyp exists at screening */
       else
         qpol = pol->next;
     } /* pass to find any detected polyp-cancer */
     prob = genscrn();
     if (found==YES) {
      if (prob > DECLINE)
         result = TRUEPOS;
      else
         result = FALSENEG;
     }
     else if (any==YES)
      result = FALSENEG;
     else {
      if (prob <= SPECFOBT) {
         prob = genscrn();
         if (prob > DECLINE)
            result = FALSEPOS;
         else
            result = TRUENEG;
      }
      else
         result = TRUENEG;
     }
     return result;
}
```

Figure A6: C code which applies FOBT screen *(continued)*

```
/*****************************************************************

    fobt_outcome

    routine to record the outcome of the screening

    inputs
    p - pointer to screen chain member
    ptr - pointer to population unit
    age - age at screen
    result - result of screen

    calls routines
    detect_stage
    detect_cancer

    returns
    int continue (YES or NO)

*****************************************************************/
int fobt_outcome( struct screen *p,
              struct member *ptr,
              double age,
              int result)
{
     int contin;
     int stage;
     struct polyp *pol;

     switch (result) {

     case TRUEPOS:
      p->colonosc = 1;
      p->cclin = 1;
      if (genscrn() <= PERFRATE)
         p->perf = 1;
      ptr->age_detect = age;
      for(pol=ptr->ppolyp; pol!=NULL; pol=pol->next) {
          stage = detect_stage(pol, age);
```

Figure A7: C code which records FOBT screen outcome

```
          if (stage == SMPOLYP) {
             cancel_polyp(pol, ptr);
          } /* cancel small polyps (which won't be followed up) */
          if (stage > ptr->stage_detect) {
              ptr->stage_detect = stage;
          } /* determine stage at polyp/cancer detection */
      } /* traverse polyp chain */
      if (ptr->stage_detect > ptr->stage_surface)
          ptr->stage_detect = ptr->stage_surface;
      if (ptr->stage_detect == ptr->stage_surface) {
          ptr->detect_death = ptr->canc_death;
          ptr->screen_cure = ptr->canc_cure;
      }
      else
          ptr->detect_death = detect_cancer(ptr, age);
      if (ptr->stage_detect==SMPOLYP)
         contin = YES;
      else
         contin = NO;
      break; /* true positve loop */
     case FALSEPOS:
      p->colonosc = 1;
      p->cclin = 1;
      if (genscrn() <= PERFRATE)
         p->perf = 1;
      contin = YES;
      break; /* false positve loop */
     case TRUENEG:
      contin = YES;
      break; /* true negative loop */
     case FALSENEG:
      contin = YES;
      break; /* false negative loop */
     } /* switch screening outcome */

     return contin; /* returns whether or not to continue screening
*/
}
```

Figure A7: C code which records FOBT screen outcome ***(continued)***

Appendix B. The pseudo-random number generator

The heart of any microsimulation modelling is the efficient generation of pseudo-random numbers. So an important pre-cursor to developing the model is to find a good pseudo-random number generator.

Most computer systems incorporate a pseudo-random number generator, but there are drawbacks in relying on any of these. Firstly I need to be confident of the properties of the generator. Most modern computer systems may have reliable generators, but there are enough examples of defective generators reported in the literature for me not to rely on them (Schrage 1979). Secondly, the model should produce results that are the same, where the inputs are the same, whichever computer is being used. So I decided to use a machine independent, portable pseudo-random number generator.

Ripley (1987) recommends the generator developed by Wichmann and Hill (1982). This is an efficient and portable pseudo-random number generator whose properties have been widely studied (Wichmann & Hill 1983, Zeisel 1986, McLeod 1985). It has a period exceeding 6.95×10^{12}, which is easily enough for my purposes (Wichmann & Hill 1983), and it passes a variety of statistical tests of randomness (Wichmann & Hill 1982). It is short, fast and easily programmed in most computer languages. It was written originally in FORTRAN but is easily translated into the C language.

The C code for this generator is listed in figure B.1. The final model has two such generating functions—`genpopn()`, which is listed in the figure, and `genscrn()` which is identical except for the name. The reason for this is that the particular characteristics of the simulated population (such as, for example, ages at death and distribution of cancer cases) depend on the particular stream of random numbers generated in a given run of the program. The function `genpopn()` was used to generate the stream of numbers to define the population characteristics and the function `genscrn()` was used to generate the numbers which controlled screening participation and outcomes. This ensured that different screening programs were applied to the same population, so that differences between screening outcomes were not confounded with random variations in the characteristics of the population.

Figure

```
#define seed1 10000
#define seed2 20000
#define seed3 30000

double genpopn (void)
{
     static int ix = seed1;
     static int iy = seed2;
     static int iz = seed3;
     double random, hold1;
     int hold2;

     ix = 171*(ix%177) - 2*(ix/177);
     iy = 172*(iy%176) - 35*(iy/176);
     iz = 170*(iz%178) - 63*(iy/178);

     if (ix < 0)
       ix += 30269;
     if (iy < 0)
       iy += 30307;
     if (iz < 0)
       iz += 30323;

     hold1 = (double)ix/30269.0 + (double)iy/30307.0  +
(double)iz/30323.0;
     hold2 = hold1;
     random = hold1 - (double)hold2;

     return (random);
}
```

Figure B1: The C code implementation of the Wichmann and Hill pseudo random number generator